Institut für Konstruktionstechnik
und Technisches Design
Universität Stuttgart
Prof. Dr.-Ing. Thomas Maier

Studienarbeit

cand. mach. Philipp Jordan

Auswahl einer geeigneten Methode zur Usability Evaluation

Institut für Konstruktionstechnik
und Technisches Design
Universität Stuttgart
Prof. Dr.-Ing. Thomas Maier

Studienarbeit

cand. mach. Philipp Jordan

Matrikelnummer 216 83 64

im Hauptfach

Konstruktionstechnik

Thema:

Auswahl einer geeigneten Methode zur Usability Evaluation

Diese theoretische Studienarbeit ist im engen Kontakt mit dem Institut auszuarbeiten und bleibt dessen Eigentum.

Ausgabetag: 01.02.2008 **Abgabetag: 25.06.2008**

Hauptfachprofessor: Prof. Dr.-Ing. T. Maier

Betreuer: Dr.-Ing. M. Schmid

Aufgabenstellung

Inhaltsverzeichnis

Seite

1 Einleitung 6

2 Begriffsklärung 7

2.1 Usability – Gebrauchstauglichkeit 7

2.2 Evaluation 7

2.3 Test 8

2.4 Interface 8

2.5 Interaktives System 8

2.6 Usability-Problem 9

2.7 Usability-Experte 9

3 Einführung in die Normen 11

3.1 DIN 9241-11 12

3.2 DIN 9241-110 14

3.3 DIN 14915-1 17

3.4 DIN 13407 18

3.5 Bildschirmarbeitsverordnung 19

3.6 Zusammenfassung der vorgestellten Normen und Gesetze 20

4 Methoden der Usability-Evaluation 21

4.1 Hierarchische Gliederung der Methoden zur Usability-Evaluation 22

4.2 Benutzer-Test 23

4.2.1 Fragestellungen eines Testplans vor einem Benutzertest 23

4.2.2 Das stationäre Usability-Lab 24

4.2.3 Das mobile Usability-Lab 28

4.2.4 Detaillierter Ablauf des Verfahrens 29

4.2.5 Mögliche Erhebungsmethoden beim Benutzertest 31

4.2.6 Methode des lauten Denkens/Think-Aloud-Protocol 32

4.2.7 Videoanalyse/Videofeedback 34

4.2.8 Eingabeprotokolle/Logfiles 35

4.2.9 Aufmerksamkeitsanalyse/Eye-Tracking 36

4.2.10 Fragebögen und standardisierte Interviews 39

4.2.11 Observation 47

4.3 Heuristische Evaluation 48

4.3.1 Heuristiken 48
4.3.2 Detaillierter Ablauf des Verfahrens 51
4.3.3 Optimale Anzahl an Experten 54
4.3.4 Abwandlungen und Modifikationen 55
4.4 Cognitive Walkthrough 59
4.4.1 Theorie 60
4.4.2 Detaillierter Ablauf des Verfahrens 63
4.4.3 Dokumentation des CWT 67
4.4.4 Lösungsfindungsprozess/Revision des Interfaces 67
4.4.5 Streamlined Cognitive Walkthrough 68
4.5 KLM-GOMS 69
4.5.1 Die vier Komponenten des GOMS-Modells 69
4.5.2 Einsatzbereich des KLM-GOMS Modells 71

5 Bewertung 73
5.1 Bewertungskriterien 74
5.1.1 Kriterien mit praktischer Relevanz 74
5.1.2 Kriterien mit wissenschaftlicher Relevanz 75
5.2 Bewertungstabelle 76

6 Theoretische Beschreibung der Methoden am einem Beispiel 79
6.1 Produktdefinition - Zielgruppendefinition 79
6.1.1 Definition der Ausgangssituation vor der Evaluation 80
6.1.2 Definition des Evaluationsziels 80
6.1.3 Ableitung von Arbeitsaufgaben aus den Kernfunktionen 80
6.1.4 Formulierung von Aufgabenszenarien 82
6.1.5 Zielgruppendefinition 83
6.2 Durchführung des CWT 85
6.3 Durchführung der Heuristischen Evaluation 90
6.4 Durchführung eines Benutzertests 93
6.5 Vorteile und Nachteile der verschiedenen Methoden 96

7 Fazit und Kommentar 98

8 Quellenverzeichnisse 100
8.1 Literatur- und Quellenverzeichnis 100
8.2 Abbildungsverzeichnis 105
8.3 Tabellenverzeichnis 106

1 Einleitung

"Es muss wunderbar sein; ich habe keine Ahnung, was es bedeutet." (Jean Baptiste Molière 1622-1673)

Technische Produkte und Gebrauchsgegenstände haben den Alltag des Menschen erobert und sind heutzutage ein fester Bestandteil der modernen Gesellschaft geworden.

Sie dienen der Unterstützung bei bestimmten Arbeitsprozessen und zur Erleichterung bei der Erledigung bestimmter Arbeitsaufgaben. Fahrkarten- und Bankautomaten sind Beispiele für die alltägliche Konfrontation des Menschen mit technischen Produkten und interaktiven Systemen.

Im Privatbereich haben sich Mobiltelefone, MP3-Player, Stereoanlagen und sonstige Kommunikations- und Entertainmentprodukte durchgesetzt. Der Trend der Hersteller zur Multifunktionalisierung solcher Geräte führt zu immer komplexeren Interfacegestaltungen und erhöhten Anforderungen an die zukünftigen Bediener.

Die Berücksichtigung einer intuitiven, durchdachten Gestaltung des Benutzerinterfaces eines technischen Produkts kann über dessen Erfolg oder Versagen am Markt entscheiden. Die Reduktion der Gebrauchstauglichkeit eines Produktes auf seine reine Funktionalität ist im heutigen Wettbewerb nicht mehr zeitgemäß. Moderne technische Produkte und interaktive Systeme müssen in diesem Sinne nicht nur funktionieren, sondern auch den Benutzer und dessen individuelle Ziele und Aufgaben berücksichtigen.

Diese neuen Anforderungen bei der Gestaltung der Mensch-Maschine-Schnittstelle eines technischen Produktes werden durch den Begriff der Gebrauchstauglichkeit, der Usability, zusammengefasst und beschrieben.

Es gibt verschiedene wissenschaftliche und praktische Ansätze und Methoden, um die Usability von technischen Produkten und interaktiven Interfaces zu evaluieren. Gemeinsames Ziel all dieser Methoden ist die Bewertung der Bedienbarkeit des technischen Produkts.

Im Rahmen dieser Studienarbeit werden verschiedene Verfahren und Methoden zur Evaluation der Gebrauchstauglichkeit von technischen Produkten und Systemen vorgestellt. Anhand einer Bewertung auf Basis einer Literaturrecherche werden die Vor- und Nachteile der Methoden aufgezeigt, deren theoretische Durchführung beschrieben und abschließend anhand eines fiktiven Anwendungsbeispiels exemplarisch erläutert.

2 Begriffsklärung

Im folgenden Kapitel werden die wichtigsten Begriffe zum Thema Usability-Evaluation vorgestellt und anhand relevanter Normen und Literatur definiert.

2.1 Usability – Gebrauchstauglichkeit

Die Gebrauchstauglichkeit eines Gutes ist dessen Eignung für seinen bestimmungsgemäßen Verwendungszweck, die auf objektiv und nichtobjektiv feststellbaren Gebrauchseigenschaften beruht und deren Beurteilung sich aus individuellen Bedürfnissen ableitet /4/.

Weltweit anerkannt als die maßgebliche Norm für die Gestaltung von Systemen mit einer hohen Ergonomie und Gebrauchstauglichkeit, definiert die DIN EN ISO 9241-11 Gebrauchstauglichkeit, als das Ausmaß in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen /7/. In Kapitel 3.1. wird diese Norm als Grundlagennorm vorgestellt und erläutert.

2.2 Evaluation

Unter dem Begriff Evaluation versteht man die Bewertung eines Prüfergebnisses für ein interaktives System oder eines seiner Merkmale hinsichtlich seiner Wirkungen auf die Benutzertätigkeit (Effizienz), die Benutzereinstellung (Zufriedenstellung) oder das Ergebnis der Nutzung (Effektivität) /3/.

Nach Sarodnick /32/ bezeichnet der Begriff Evaluation allgemein eine systematische und möglichst objektive Bewertung eines geplanten, laufenden oder abgeschlossenen Projekts, dessen Ziel die Beantwortung von spezifischen Fragestellungen und/oder die Erhebung eines Zielerreichungsgrades ist. Mögliche Evaluationskriterien können zum Beispiel die in Kapitel 3.1 eingeführten Maße der Gebrauchstauglichkeit, also Effektivität, Effizienz und Zufriedenheit sein.

Eine weitere Definition findet sich bei Nielsen und Mack /25/. Sie gliedern den Begriff in Usability Inspection und Usability Testing:

"*Usability inspection is the generic name for a set of methods based on having evaluators inspect or examine usability-related aspects of a user interface.*"

"*User testing with real users is the most fundamental usability method and is in some sense irreplaceable, since it provides direct information about how people use computers and what their exact problems are with the concrete interface being tested.*"

Unter Usability-Inspektionen werden demnach alle Verfahren verstanden, die statt echten Probanden Experten als Evaluatoren einsetzen. Usability-Tests bezeichnen demnach Verfahren die echte Nutzer zur Datenerhebung einsetzen.

Diese erste Differenzierung der Evaluation in formal-analytische Expertenmethoden und empirisch-aufgabenorientierte Testmethoden wird im Rahmen dieser Arbeit beibehalten.

2.3 Test

Als Test bezeichnet man den Vorgang einer Untersuchung bzw. die Gesamtheit, der zur Durchführung notwendigen Requisiten und jede Untersuchung, sofern sie Stichprobencharakter hat /1/.

2.4 Interface

Unter einem Interface versteht man die Art und Anzahl der Stellteile und Anzeigen und deren Anordnung auf einer Tragwerksgestalt /19/.

Synonyme Begriffe für das Interface sind:

- Mensch-Maschine-Schnittstelle
- Mensch-Maschine-Interface
- Mensch-Maschine-Interaktion
- Mensch-Maschine-Kommunikation

Das Interface ist die Kontaktstelle zwischen dem Benutzer und dem interaktiven System und dient der Interaktion und dem Dialog zwischen dem Menschen und dem technischen System. Die Interaktion kann über alle Sinne des Menschen erfolgen sowohl klassisch mittels Kombination von Monitor mit Tastatur, Maus oder Mikrophon- und Lautsprecher als auch in kinetischer, also mechanischer Form, zum Beispiel, bei der Gangschaltung in einem Auto. Je anwenderfreundlicher diese Schnittstelle konzipiert ist, desto leichter ist der Benutzer in der Lage damit umgehen.

2.5 Interaktives System

Ein Interaktives System ist die Kombination von Hardware- und Softwarekomponenten, die Eingaben von einem Benutzer empfangen und Ausgaben zu einem Benutzer übermitteln, um ihn bei der Ausführung einer Arbeitsaufgabe zu unterstützen. Oft wird statt „interaktives System“ nur der Begriff „System“ verwendet /5/.

2.6 Usability-Problem

Unter Berücksichtigung des Nutzungskontexts eines technischen Systems und des Erfahrungsgrades der potentiellen Nutzer werden Usability-Probleme folgendermaßen definiert:

Ein Usability-Problem liegt vor, wenn Aspekte eines Systems es Nutzern mit hinreichender Domänenerfahrung unangenehm, ineffizient, beschwerlich oder unmöglich machen in einem typischen Anwendungskontext die Ziele zu erreichen, für deren Erreichung das System erstellt wurde /33/.

2.7 Usability-Experte

Die Merkmale eines Usability-Experten sind weder verbindlich definiert, noch genormt. Im Leitfaden Usability der Deutschen Akkreditierungsstelle für Technik /3/ werden jedoch Kompetenzen aufgelistet, die ein Usability-Experte besitzen sollte:

Kompetenzbereich eines Usability-Experten	**Beispiel**
Analysekompetenz	Auf der Basis von Nutzungskontexten und Erfordernissen Nutzungsanforderungen kontextbezogen und lösungsneutral spezifizieren zu können.
Designkompetenz	Auf der Basis von Kernaufgaben Aufgabenmodelle konstruieren und hierfür Nutzungsszenarien entwerfen zu können.
Prüfungskompetenz	Teilnehmende Beobachtungen mit Nutzern, planen, durchführen und dokumentieren zu können.
Bewertungskompetenz	Die Auswirkung von erkannten Abweichungen von Prinzipien, Gestaltungsregeln, Nutzungsanforderungen und Prüfkriterien erkennen und bewerten zu können.
Soziale Kompetenz	Kontextbezogene Interviews mit prospektiven Nutzern durchführen zu können.
Technische Kompetenz	Mutmaßliche technische Nicht-Machbarkeiten einschätzen zu können.

Tabelle 2.1: Kompetenzen eines Usability-Experten

Ein Usability-Experte zeichnet sich durch Kenntnis der verwendeten Evaluationsmethodik und der ihr zugrunde liegenden Theorie, sowie praktischer Erfahrungen im Bereich der nutzerzentrierten Interfacegestaltung und Bewertung technischer Produkte und interaktiver Systeme aus. Dieser Erfahrungsgrad eines Experten wird als Usability-Expertise bezeichnet.

Die Domänenexpertise eines Experten bezeichnet die Erfahrung und das Hintergrundwissen in Bezug auf das zu bewertende Produkt oder interaktive System. In hochkomplexen Anwendungsdomänen wie zum Beispiel der digitalen Produktionsplanung, innerhalb derer nicht nur das zu bewertende System, sondern auch die damit verketteten Prozesse berücksichtigt werden müssen, kann es für Usability-Experten problematisch sein, in einem begrenzten Zeitraum sich das benötigte Hintergrundwissen zur Bewertung anzueignen.

Daher ist in hochkomplexen Anwendungsdomänen die Verwendung von Experten mit Usability- sowie Domänenexpertise, sogenannter Doppelexperten, entscheidend für die Qualität des Ergebnisses der Evaluation.

3 Einführung in die Normen

In diesem Kapitel werden relevante Normen und Gesetze zum Bereich Software-Ergonomie und Usability-Engineering vorgestellt. Diese Normen beschreiben detailliert die zu verwendenden Fachtermini und stellen einen Konsens über gebündeltes Fachwissen im Bereich Software-Ergonomie und Usability-Evaluationen dar.

Als die maßgebliche Normenreihe, für die Gestaltung von interaktiven Produkten und Systemen mit einer hohen Gebrauchstauglichkeit, wird die DIN EN ISO 9241 ausführlich erläutert. Diese Normenreihe definiert wichtige Fachbegriffe und Termini zum Thema Gebrauchstauglichkeit. Des Weiteren werden zusätzliche Normen und Gesetze, wie die Bildschirmarbeitsverordnung, die für die ergonomische Gestaltung und Bewertung von interaktiven Systemen relevant sind, kurz vorgestellt und erläutert.

Die hier vorgestellten Normen und Richtlinien werden traditionell dem Bereich der Software-Ergonomie von Bildschirmsystemen zugeordnet. Durch die immer größer werdende Verbreitung von Computersystemen und Softwareprogrammen als Massenprodukte, den Siegeszug des Internets, sowie die Einführung der graphischen Benutzeroberflächen in den späteren 1980er und frühen 1990er Jahren wurden Computer und Bildschirmsysteme zunehmend im geschäftlichen sowie privatem Bereich Standard.

Gleichzeitig entstand jedoch Handlungsbedarf für die Menschengerechte Gestaltung von Bildschirmsystemen. Computer und interaktive Systeme sollten Menschen ihre Arbeit nicht nur ermöglichen, sondern sie auch bei ihrer Arbeit angemessen unterstützen, Hilfe anbieten und vor allem Fehler vermeiden.

Diese Erkenntnis ist aktuell wichtiger denn je, da heute in nahezu jedem komplexeren Produkt ein elektronischer Mikroprozessor verbaut und ein Interface vorhanden ist. Die Liste der technischen Geräte, mit denen der Mensch heutzutage im Alltag konfrontiert wird, ist lang. Die Kiosksysteme (auch Walk-Up-and-Use Systeme genannt), also jegliche Art von Automaten wie Bankautomaten, Parkscheinautomaten, Kaffeeautomaten und Ticketautomaten sind heutzutage allgemein akzeptiert und weit verbreitet. Tragbare multifunktionale Systeme wie Mobiltelefone, Laptops und Digitalkameras sowie Heimgeräte wie multifunktionale Festplattenrecorder, Digitalreceiver und auch Fernseher vereinen immer mehr Funktionen in einem Gerät und werden in ihrer Bedienung meistens immer komplexer.

Diese Normen, als Regelwerke und Leitlinien für die Menschengerechte Gestaltung von Bildschirmsystemen und Arbeitsplätzen können dazu benutzt werden, den schmalen Grat zwischen der Menschen- und Aufgabengerechten Gestaltung eines interaktiven Systems oder technischen Produkts und der zunehmenden Funktionsvielfalt und Komplexisierung ihrer Produktgestalt, erfolgreich zu durchschreiten. Sie sind daher auf nahezu jedes interaktive System mit dem Benutzer interagieren anwendbar.

3.1 DIN 9241-11

Gebrauchstauglichkeit (englisch: Usability) wird in der DIN EN ISO 9241: Teil 11 definiert als das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen /7/.

In diesem Zusammenhang umfasst der Begriff Produkt, das technische System das evaluiert werden soll und mit dem der Benutzer interagiert. Dabei kann es sich um ein simples Werkzeug bis hin zur hochkomplexen, interaktiven Softwareumgebung handeln. Die Definition beschränkt sich keinesfalls nur auf Bürotätigkeiten mit Bildschirmgeräten, sondern ist universell einsetzbar, wenn Benutzer mit einem Produkt arbeiten um ein bestimmtes Ziel zu erreichen.

Effektivität beschreibt die Genauigkeit und Vollständigkeit, mit der ein Benutzer ein bestimmtes Ziel erreichen kann, Effizienz ist der dazu eingesetzte Aufwand. Effektivität lässt sich beispielsweise durch einen Zielerreichungsgrad der vorhandenen Aufgabenstellung des Benutzers quantifizieren. Effizienz wäre in diesem Zusammenhang dann die Anzahl der benötigten Schritte bzw. die Zeit, die der Benutzer für die Erledigung der Aufgabe benötigt. Sie setzt den Grad der Effektivität ins Verhältnis zum Aufwand an Ressourcen.

Als zufriedenstellend werden die subjektiven Erfahrungen des Benutzers mit dem Produkt verstanden. Dazu zählt beispielsweise die Freiheit von Beeinträchtigungen bei der Bedienung, Wertschätzung des Produkts an sich oder auch ein hoher Grad an Effizienz. Eine hohe freiwillige Benutzungsfrequenz des Produkts und eine geringe Anzahl an Probleme und Beschwerden bei der Benutzung wären mögliche Maße der Zufriedenstellung.

In der Norm wird hervorgehoben, dass die Gebrauchstauglichkeit eines Produktes vom Nutzungskontext abhängt, und dass die besonderen Umstände unter denen das Produkt benutzt wird, den Grad der Gebrauchstauglichkeit beeinflussen. Der Nutzungskontext des Produkts beinhaltet den Benutzer, die Arbeitsaufgabe, die Arbeitsmittel (Hardware, Software und Materialien) sowie die physische und soziale Umgebung.

Die Zusammenhänge der einzelnen Definitionen lassen sich anschaulich im Anwendungsrahmen für Gebrauchstauglichkeit, Bild 3.1, darstellen. Der Benutzer bildet mit der Arbeitsaufgabe, den Arbeitsmitteln und der Umgebung den Nutzungskontext und interagiert mit dem Produkt um sein angestrebtes Ziel zu erreichen. Über die eingeführten Maße der Gebrauchstauglichkeit lässt sich nach der Benutzung, die Usability des Produkts in dem vorgegebenen Kontext qualitativ und quantitativ bestimmen.

Usability beschreibt also die Qualität eines technischen Systems in einem definierten Nutzungskontext. Die systematische und methodische Erzeugung von Usability nennt

man entsprechend Usability-Engineering. In diesem Zusammenhang ist Usability ein Teilziel der Ergonomie, also der Anpassung der Technik an den Menschen.

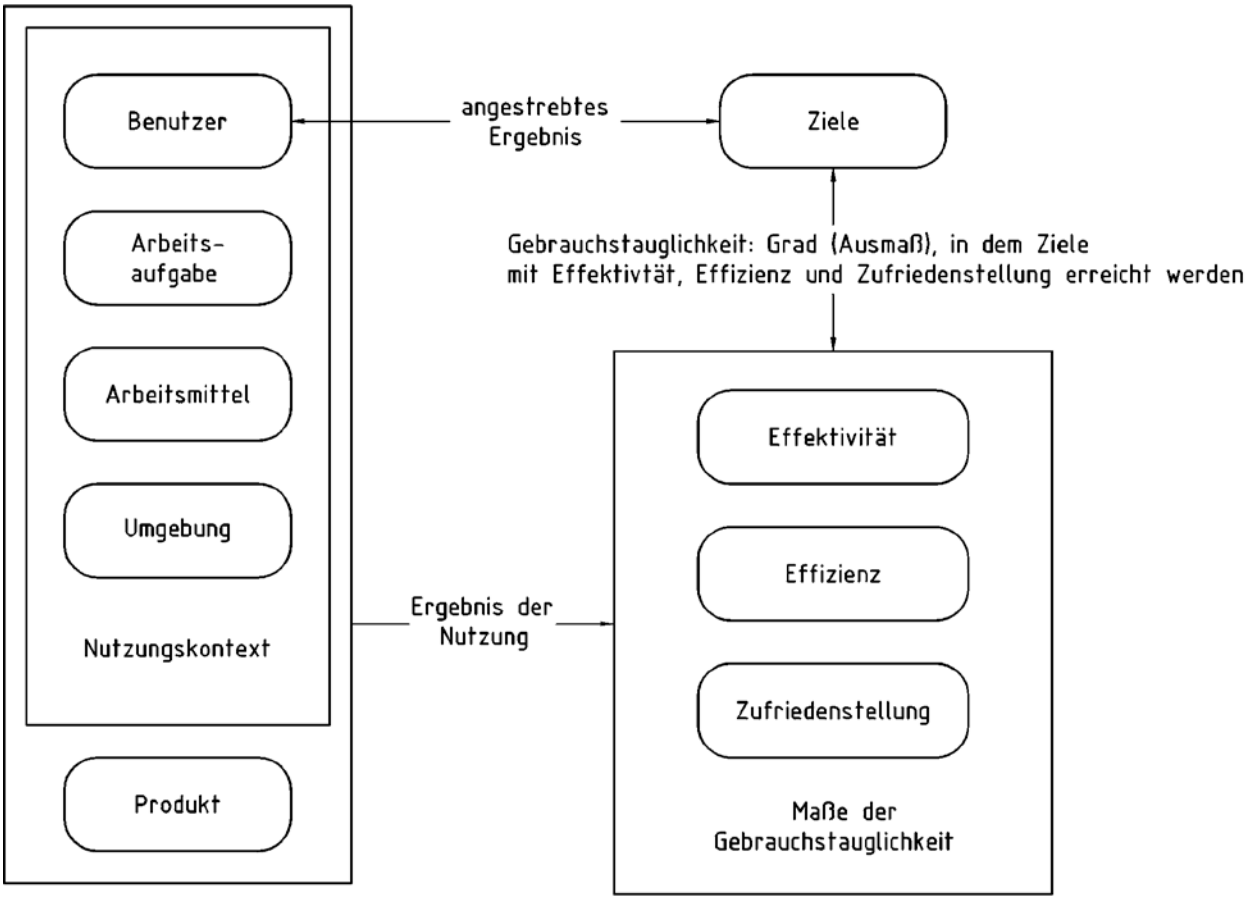

Bild 3.1: Anwendungsrahmen für Gebrauchstauglichkeit /50/

3.2 DIN 9241-110

In der DIN 9241-110 werden sieben Grundsätze der Dialoggestaltung /8/ für die ergonomische Gestaltung von interaktiven Systemen formuliert. Diese Grundsätze sind sehr allgemein formuliert und daher auf nahezu jedes interaktive System anwendbar und werden im Folgenden kurz vorgestellt:

1. **Aufgabenangemessenheit**

Die Aufgabenangemessenheit eines interaktiven Systems beschreibt das Ausmaß, in dem es einen Benutzer bei Erledigung seiner Arbeitsaufgabe unterstützt, was bedeutet, dass die Funktionalität und eingesetzte Dialogtechnik auf den charakteristischen Eigenschaften der Arbeitsaufgabe basiert, anstatt auf der zur Aufgabenerledigung eingesetzten Technologie.

Ein Beispiel für Aufgabenangemessenheit eines technischen Systems, ist die Voreinstellung des Abfahrtsbahnhofs eines Fahrkartenautomaten auf die Station, an der das Ticket gelöst wird.

2. **Selbstbeschreibungsfähigkeit**

Interaktionen zwischen Mensch und Maschine sollen selbsterklärend und transparent gestaltet sein, woraus folgt, dass für den Nutzer zu jedem Zeitpunkt ersichtlich ist, an welcher Stelle er sich in einem speziellen Dialog befindet und was die nächstmöglichen Schritte innerhalb dieses Dialoges sind.

Ein Beispiel für die Selbstbeschreibungsfähigkeit eines technischen Systems, ist demnach die eindeutige und unmissverständliche Beschriftung der Tasten eines Telefons für spezielle Funktionen.

3. **Erwartungskonformität**

Unter der Erwartungskonformität eines Dialoges wird die Beachtung und Berücksichtigung von allgemeinen Konventionen im jeweiligen Anwendungsbereich und Nutzungskontext des zu erwartenden Benutzers verstanden.

Ein gutes Beispiel für die Erwartungskonformität ist die Übernahme von Begriffen und Wörtern aus dem realen Leben in das interaktive System. Interaktive Systeme, wie etwa eine Software zur Abwicklung elektronischer Bankgeschäfte, sollten demnach gängige Begriffe wie Girokonto und Überweisung als Terminologie aus dem realen Bankengeschäft übernehmen.

4. **Lernförderlichkeit**

Die Lernförderlichkeit beschreibt das Ausmaß, in dem ein lernender Benutzer innerhalb eines interaktiven Systems durch selbiges unterstützt und angeleitet wird. Hier wird unter anderem die Forderung nach einem minimalen Lernaufwand bei Benutzung einer Software gestellt.

Als Beispiel für die Lernförderlichkeit soll hier ein Archivierungsprogramm betrachtet werden, das vor der erstmaligen Nutzung dem Benutzer das grundlegende Konzept eines Archivs innerhalb einer Einführung erläutert.

5. **Steuerbarkeit**

Ist ein Benutzer in der Lage in einem Dialog innerhalb eines interaktiven Systemen, Ablauf, Geschwindigkeit und Richtung des Dialogablaufs zu bestimmen, so ist dieser Dialog steuerbar.

Hier dient als Anschauungsbeispiel die Steuerbarkeit eines Mobiltelefons, welches dem Benutzer erlaubt einen ankommenden Anruf anzunehmen, abzuweisen oder direkt im Telefonbuch zu speichern.

6. **Fehlertoleranz**

Eine hohe Fehlertoleranz bedeutet, dass falsche Eingaben innerhalb eines Dialoges in interaktiven Systemen mit keinem oder minimalem Aufwand korrigiert werden können.

Beispielsweise wird in einen Online-Formular auf Pflichtfelder beim Ausfüllen hingewiesen oder es werden Hinweise zur Fehlerkorrektur gegeben.

7. **Individualisierbarkeit**

Unter der Individualisierbarkeit eines Dialoges versteht man die Möglichkeit des Benutzers, die Darstellung und Anordnung der Informationen während der Mensch-System-Interaktion auf seine persönlichen Bedürfnisse hin individuell zu gestalten.

Die Auswahlmöglichkeit der Sprache an einem beliebigen Kiosksystem, wie einem Bankautomaten, ist ein sehr einfaches und anschauliches Beispiel für die Individualisierbarkeit eines Dialoges.

Diese Grundsätze können für die Analyse, Gestaltung und Bewertung von Dialogsystemen benutzt werden, jedoch unter Berücksichtigung einer individuellen Gewichtung, die sich nach den Merkmalen des Benutzers, den Arbeitsaufgaben, der Arbeitsumgebung und der Dialogtechnik richtet.

In der Norm wird betont, dass der Benutzer die Anforderungen an die Ergonomie des Systems stellt. Durch die Anwendung der Leitsätze können Nutzungsprobleme wie irreführende Informationen oder unnötige, überflüssige Schritte, vermieden werden. Die

Grundsätze stellen allgemeine Zielsetzungen bei der Gestaltung interaktiver Systeme dar, die zu gebrauchstauglichen System führen sollen.

Der Problematik der Interferenz der verschiedenen Gestaltungsanforderungen beim praktischen Umsetzen in interaktiven Systemen und Produkten ist durch eine entsprechende Priorisierung Rechnung zu tragen. So gilt es abzuwägen, ob beispielsweise zugunsten einer vermeintlich hohen Fehlertoleranz eine geringe Steuerbarkeit und Individualisierbarkeit „in Kauf“ genommen wird.

Bild 3.2 zeigt die Zusammenhänge der verschiedenen Normteile der 9241 auf. Es wird hierbei deutlich, das Usability, als Qualitätsmerkmal eines technischen Systems oder interaktiven Produkts als ein ganzheitliches Konzept aufgefasst wird, das verschiedene Anforderungen wie die vorgestellten Grundsätze der Dialoggestaltung aus dem Normenteil 110 und charakteristische Eigenschaften dargestellter Informationen aus dem Normenteil 12 zusammenfasst.

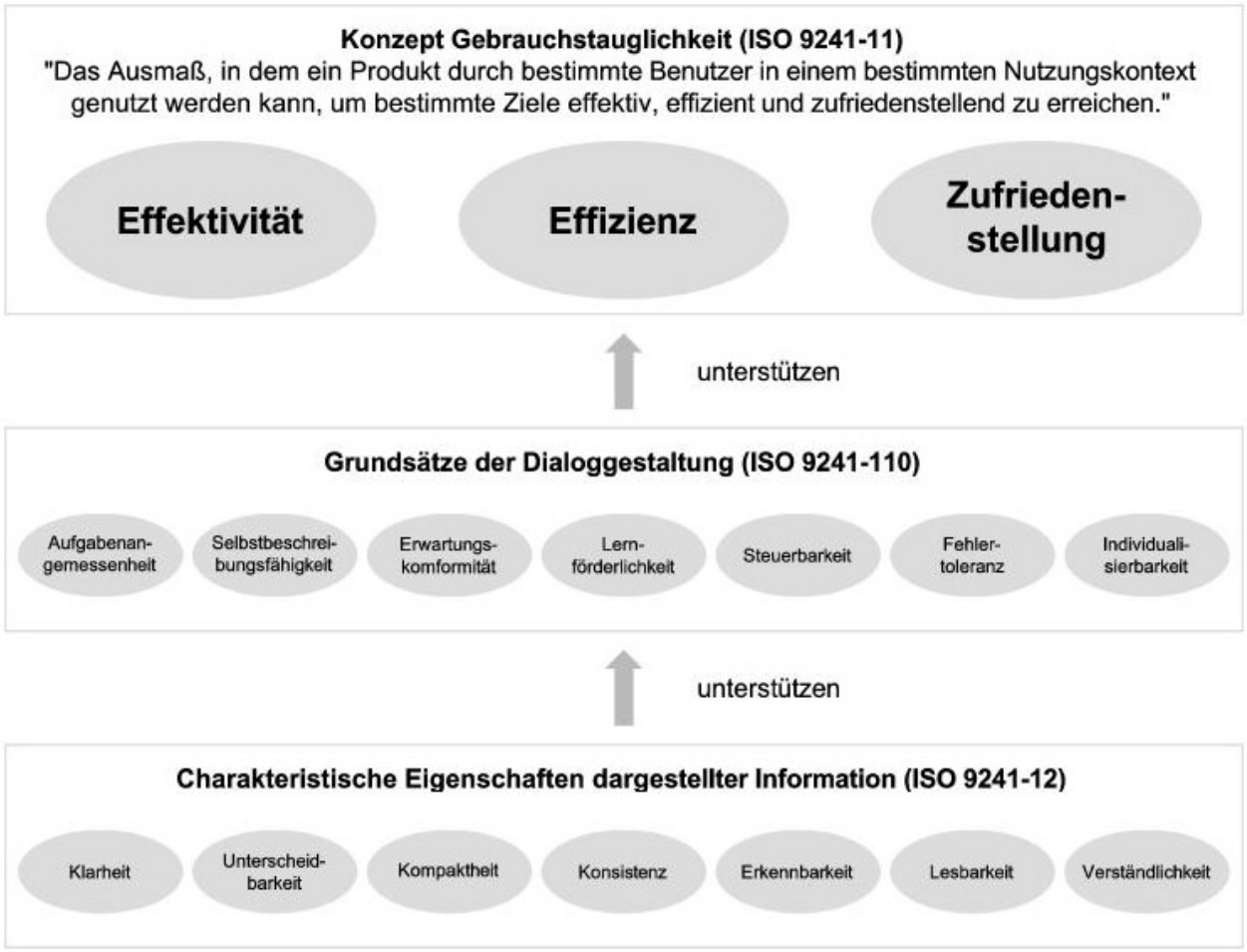

Bild 3.2: Konzept Gebrauchstauglichkeit /52/

3.3 DIN 14915-1

Die DIN 14915, Software-Ergonomie für Multimedia-Benutzungsschnittstellen, beruft sich auf die allgemeinen Gestaltungsgrundsätze der DIN 9241-110 und definiert zusätzliche spezielle Gestaltungsempfehlungen /6/ für die ergonomische Gestaltung von Multimedia-Benutzungsschnittstellen:

1. **Eignung für das Kommunikationsziel**

Eine Multimedia-Anwendung ist für das Kommunikationsziel geeignet, wenn sie so gestaltet ist, dass sie sowohl den Zielen des Anbieters der zu übermittelnden Information als auch dem Ziel oder der Aufgabe der Benutzer oder Empfänger dieser Information entspricht.

2. **Eignung für Wahrnehmung und Verständnis**

Eine Multimedia-Anwendung ist für Wahrnehmung und Verständnis geeignet, wenn sie so gestaltet ist, dass die zu übermittelnde Information leicht erfasst und verstanden werden kann.

Unter diesem Gestaltungsgrundsatz werden Kriterien wie Erkennbarkeit, Klarheit oder Kompaktheit der dargestellten Information der Multimedia-Schnittstelle verstanden.

3. **Eignung für Exploration**

Eine Multimedia-Anwendung ist für die Exploration geeignet, wenn sie so gestaltet ist, dass der Benutzer eine relevante oder interessante Information mit wenig oder keinem Vorwissen in Bezug auf Art, Umfang oder Struktur der Information oder der verfügbaren Funktionalität der Anwendung finden kann.

4. **Eignung für Benutzungsmotivation**

Falls es für die Arbeitsaufgabe angebracht ist, sollte eine Multimedia-Anwendung so gestaltet sein, dass sie für den Benutzer anregend ist, d. h. dass sie die Aufmerksamkeit des Benutzers auf sich zieht und ihn dazu motiviert, mit ihr zu interagieren.

3.4 DIN 13407

Die DIN 13407 definiert wichtige Grundsätze der benutzerorientierten Gestaltung und erläutert die Vorteile und Kennzeichen der Integration eines Benutzerorientierten Gestaltungsprozesses /5/. Wichtig im Rahmen der Usability-Evaluation sind innerhalb dieser Norm die zu beachtenden Merkmale bei einer Analyse des Nutzungskontexts:

Merkmale der Benutzer

- Kenntnisse, Fähigkeiten, Erfahrungen, physische Merkmale
- Nutzungshäufigkeit der Software, Übung
- Gewohnheiten, Vorlieben

Merkmale der durchzuführenden Arbeitsaufgaben

- Beschreibung der Aufgaben und Ziele
- Beschreibung der einzelnen Schritt im Arbeitsablauf
- Häufigkeit und Dauer
- Funktionsteilung zwischen Mensch und Maschine

Merkmale der Umgebung, in der das System benutzt werden soll

- Hard- und Softwareumgebung
- Arbeitsplatz und dessen Komponenten
- Umgebungsfaktoren wie Temperatur, Lärm, Luftfeuchtigkeit
- Rechtliche Umgebung wie z.B. die Bildschirmarbeitsverordnung
- Soziale und kulturelle Umgebung (Organisationsstruktur, Einstellungen, Arbeitspraktiken)

3.5 Bildschirmarbeitsverordnung

Als ein vom Gesetzgeber verfasstes Dokument, gibt die Bildschirmarbeitsverordnung /2/ verbindliche Anforderungen für Arbeitgeber hinsichtlich der Sicherheit und des Gesundheitsschutzes bei der Arbeit mit Bildschirmgeräten vor. Im Anhang des Gesetzes finden sich verschiedene Anforderungen für die Gestaltung eines Bildschirmarbeitsplatzes. Die Anforderungen teilen sich in folgende die vier Bereiche auf:

- Bildschirmgerät und Tastatur
- Sonstige Arbeitsmittel
- Arbeitsumgebung
- Zusammenwirken Mensch-Arbeitsmittel auf.

Innerhalb dieser Anforderungen müssen Bildschirmgeräte beispielsweise flimmer- und reflexionsfrei sein.

3.6 Zusammenfassung der vorgestellten Normen und Gesetze

Die hier vorgestellten Normen und die Bildschirmarbeitsverordnung werden als die maßgeblichen Regelwerke und Leitlinien für die ergonomische Gestaltung von Produkten und interaktiven Systemen erachtet. Sie dienen einer Vereinheitlichung der zu verwendenden Terminologie und schaffen einen gemeinsamen Rahmen zur ergonomischen Gestaltung und Bewertung von interaktiven Systemen und technischen Produkten.

In der Normenreihe der ISO 9241 werden die zentralen Begriffe und Konzepte für die Gestaltung, Messung und Bewertung von Produkten mit einer hohen Gebrauchstauglichkeit definiert.

Die Kernaussage der ISO 9241-11 ist demnach, dass die Gebrauchstauglichkeit eines Produktes immer in einem definierten Nutzungskontext betrachtet wird. Dieser Nutzungskontext beinhaltet den Benutzer, dessen Arbeitsaufgabe, die zur Verfügung stehenden Arbeitsmittel und die Arbeitsumgebung. Innerhalb dieses Arbeitsmodells kann man die Gebrauchstauglichkeit eines Produktes über die Maße der Effektivität, Effizienz und Zufriedenstellung feststellen und bewerten.

4 Methoden der Usability-Evaluation

Im folgenden Kapitel werden Methoden und Verfahren zur Evaluation von Produkten und interaktiven Systemen vorgestellt. Jede Methode vertritt dabei eine eigene Zielsetzung und verfolgt einen eigenen Ansatz zur Evaluation. Es werden insgesamt fünf verschiedene Methoden zur Evaluierung aufgezeigt, deren zugrunde liegendes Modell erläutert, Beispiele und Hinweise zur Durchführung genannt und deren Vor- und Nachteile im Vergleich zu den anderen Verfahren erörtert.

Die fünf Verfahren wurden unter den Gesichtspunkten der Verbreitung in der relevanten Forschungsliteratur und Anwendbarkeit in der Praxis ausgewählt. Einzig das vorgestellte KLM-GOMS-Modell zur Bewertung von Ausführungszeiten in Gestaltungsalternativen für Bildschirmsysteme (siehe Kapitel 4.5) als rein softwareergonomisches, stark spezialisiertes Verfahren hebt sich, in Bezug auf den Anwendungs- und Einsatzbereich, von den vier restlichen vorgestellten Methoden ab.

Es sei noch anzumerken, dass viele kommerzielle Anbieter „eigene“ Verfahren, die oft eine Mischung von Experteninspektionen- und Nutzertests darstellen, anbieten. Diese werden hier aufgrund der mangelnden Dokumentation in der Literatur nicht vorgestellt, obwohl sich all diese Methoden meistens auf eine der hier vorgestellten Methoden reduzieren lassen.

4.1 Hierarchische Gliederung der Methoden zur Usability-Evaluation

Bild 4.1 zeigt eine Kategorisierung der in dieser Studienarbeit vorgestellten Methoden zur Usability-Evaluation.

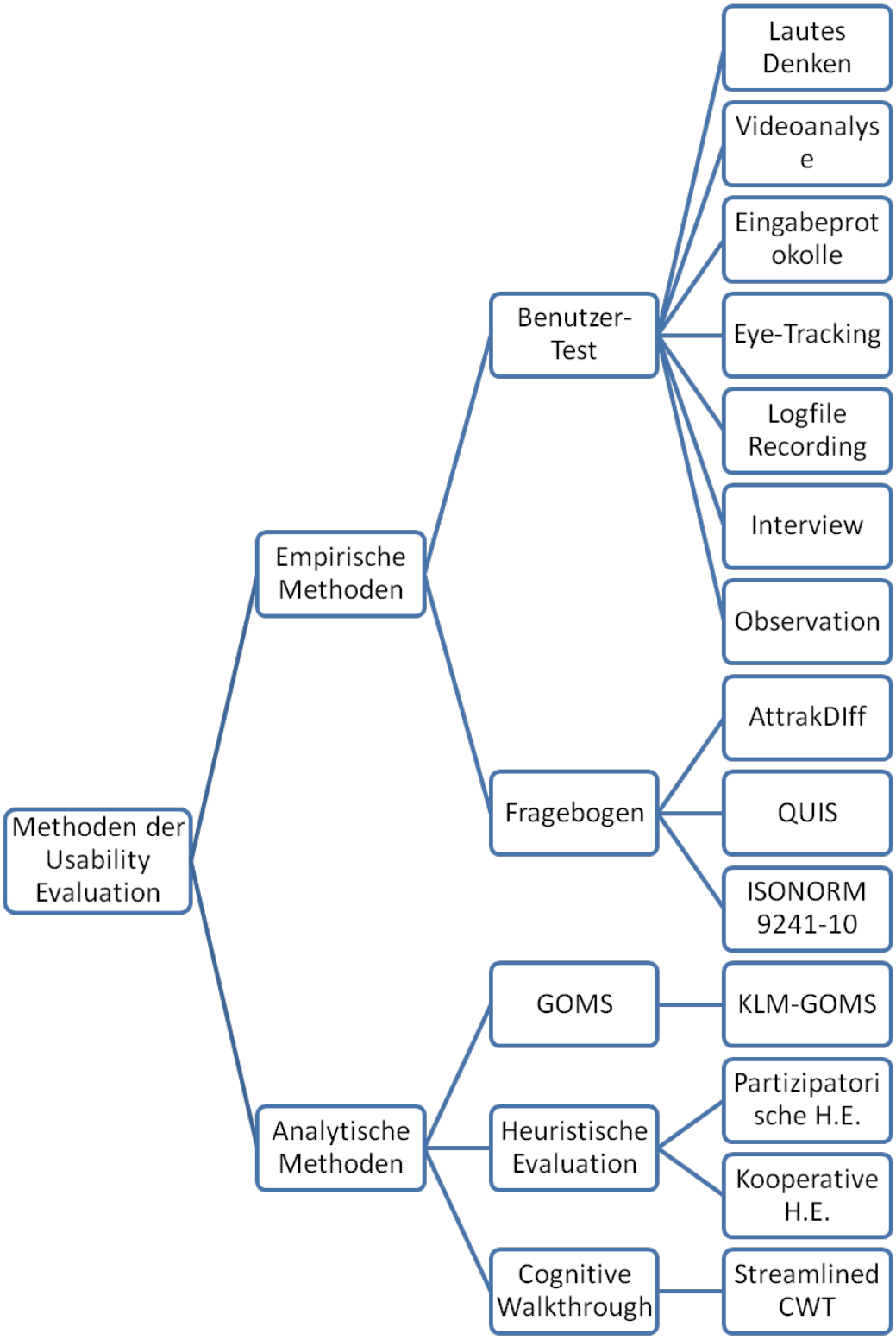

Bild 4.1: Gliederung der Methoden zur Usability-Evaluation

4.2 Benutzer-Test

Unter dem Benutzer-Test versteht man allgemein alle Verfahren und Evaluationsmethoden, die echte Nutzer zum Testen der Gebrauchstauglichkeit eines technischen Systems oder Produkts anhand von definierten Arbeitsaufgaben benutzen. Die Experten nehmen beim Benutzer-Test die Rolle der Observatoren und Moderatoren ein und versuchen über Beobachtung, Befragung und Dokumentation des Probandenverhaltens, die Usability-Probleme der Nutzer zu identifizieren.

Somit gliedern sich die Benutzer-Tests in die Kategorie der empirischen, aufgabenorientierten Verfahren ein. Die verschiedenen Testvarianten unterscheiden sich hauptsächlich darin, an welchem Ort die Evaluation stattfindet (Usability-Labor, realer Arbeitsplatz) und wie die Informationen zur Bewertung gewonnen werden (Videoanalyse, Methode des lauten Denkens, Fragebögen, Interviews).

Um bestmögliche Ergebnisse zur Informationsgewinnung zu erzielen, werden oftmals die Erhebungsmethoden kombiniert, beispielsweise die Anwendung einer Videoanalyse in Kombination mit der Methode des lauten Denkens und einem abschließenden Interview.

Im Folgenden werden die wichtigsten und bekanntesten Erhebungsmethoden bei Benutzertests angesprochen, deren Durchführung erläutert und Vor- und Nachteile der verschiedenen Erhebungsmethoden diskutiert.

4.2.1 Fragestellungen eines Testplans vor einem Benutzertest

Bevor Usability-Evaluationen durchgeführt werden, sollten die Testziele, das Anliegen sowie die Randbedingungen der Evaluation definiert werden. Um dies zu gewährleisten schlägt Nielsen daher folgende Fragestellungen /22/ für einen Benutzer-Test vor, die vor der eigentlichen Durchführung beantwortet werden sollten:

- Was sind die Ziele des Tests und welche Daten sollen erhoben werden?
- Wo und wann soll der Test stattfinden?
- Wie lange wird der Test voraussichtlich dauern?
- Bei mehreren Testterminen - Wie lange wird jeder Testdurchlauf voraussichtlich dauern?
- Welche Computersysteme sind zur Testdurchführung notwendig?
- Welche Software ist zur Testdurchführung notwendig?
- Wie ist der Status des Produkts zu Beginn des Tests?
- Wer übernimmt die Rolle des Testleiters?
- Welche Art von Probanden wird für den Test benötigt?
- Wie viele Probanden werden benötigt?
- Wo bekomme ich meine Probanden her? Wie akquiriere ich sie?

- Welche Aufgaben werden durch die Nutzer bearbeitet?
- Welche Bewertungskriterien werden verwendet um festzulegen, wann ein Proband die Testausgaben erfolgreich abgeschlossen hat?
- Welche Hilfsmöglichkeiten stehen den Probanden während des Tests zur Verfügung?
- In welchem Ausmaß hilft der Testleiter den Probanden während der Aufgabenbearbeitung?
- Welche Daten werden erhoben und wie werden die gewonnen Informationen nach der Erhebung ausgewertet?
- Was ist das Entscheidungskriterium, das aussagt, dass die Gestaltung des Produkts ein Erfolg ist?

Diese Fragenliste kann auch implizit zu der Auswahl eines Testverfahrens oder einer Erhebungsmethode benutzt werden. Wenn zum Beispiel “harte Daten“ wie Ausführungszeiten bei der Aufgabenbearbeitung gemessen werden sollen, werden durch diese Vorgabe alle Verfahren, welche keine Performance-Messung zu lassen, implizit eliminiert.

4.2.2 Das stationäre Usability-Lab

Usability-Tests können in einer definierten Testumgebung, die möglichst realistisch die echte Arbeitsumgebung simuliert, dem Usability-Labor, oder auch als “On-Site“ Tests, also direkt am Einsatzort des technischen Systems stattfinden. Dies wird durch Faktoren wie der Mobilität des Prototyps, der Verfügbarkeit von Probanden und auch die zu erwartenden Kosten determiniert.

Unter einem Usability-Labor (im folgenden Usability-Lab genannt) wird die Umgebung verstanden, in der Nutzertests zur Usability-Bewertung von interaktiven Systemen unter kontrollierten Bedingungen durchgeführt werden können. Die Ausstattung des Usability-Labs ist von den konkreten Anforderungen an Test- und Datenerhebungsverfahren abhängig.

Bild 4.2 zeigt den Aufbau eines hypothetischen Usability-Labs nach Nielsen /20/. Bild 4.3 zeigt als anschauliches Praxisbeispiel den Kontrollraum des Usability-Labs am Center for Advanced Studies and Research in Information and Communication Technologies & Society (kurz ICT&S) in Salzburg. Hier werden insgesamt drei Videokameras zur Aufzeichnung der Probandeninteraktionen verwendet.

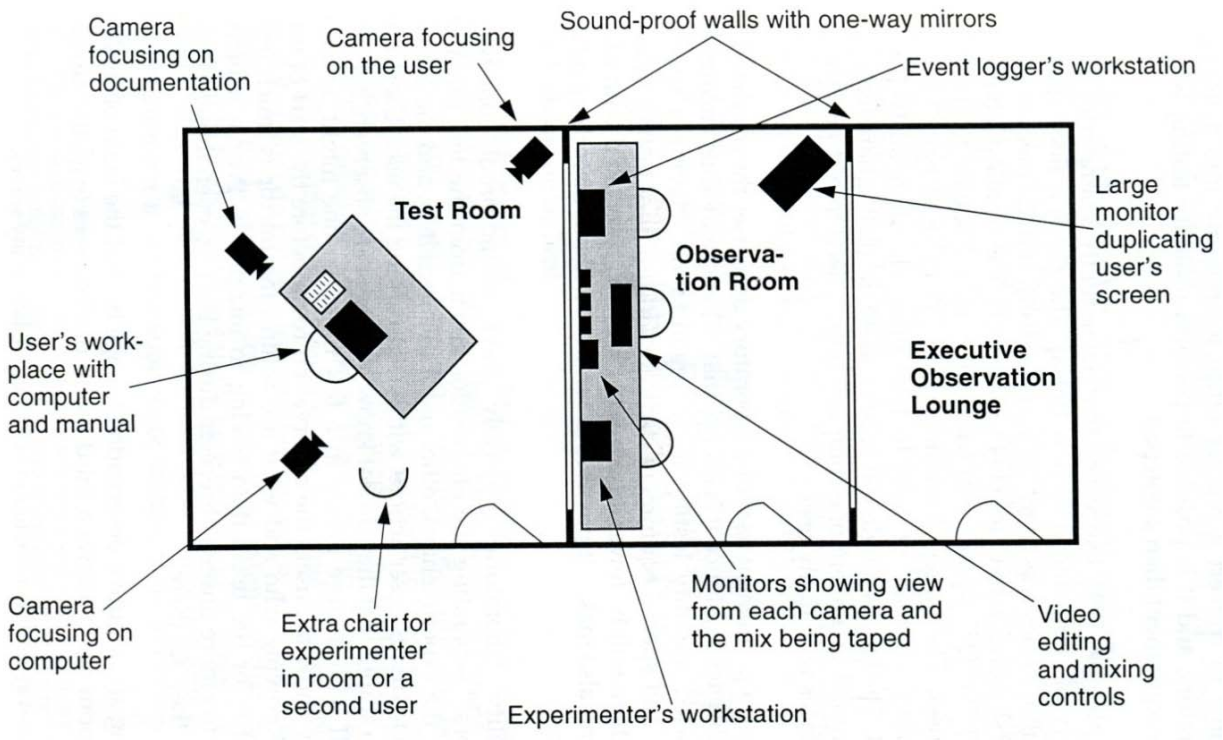
Camera focusing on documentation
Camera focusing on the user
Sound-proof walls with one-way mirrors
Event logger's workstation
Test Room
Observa-tion Room
Large monitor duplicating user's screen
User's work-place with computer and manual
Executive Observation Lounge
Camera focusing on computer
Extra chair for experimenter in room or a second user
Monitors showing view from each camera and the mix being taped
Video editing and mixing controls
Experimenter's workstation

Bild 4.2: Grundriss eines beispielhaften Aufbaus eines Usability-Labs /52/

Bild 4.3: ICTS Usability-Lab: Kontrollraum mit Blick auf Testraum /53/

Ein stationäres Usability-Lab, wie in Bild 4.2 und Bild 4.3 dargestellt, besteht aus einem Testraum und einem Kontrollraum. Im Testraum findet die Aufgabenbearbeitung durch den Probanden statt, im Kontrollraum wird der Test überwacht. Beide Räume sind durch einen Einwegspiegel voneinander getrennt. Die schalldichte Ummantelung des Testraums sorgt für eine störungsfreie Aufgabenbearbeitung durch die Probanden und erlaubt gleichzeitig, die Diskussion und Regie der Beobachter im Kontrollraum. Optional befindet sich in dem in Bild 4.2 beschriebenem Layout noch ein weiterer Beobachtungsraum, der dem Entwicklungsteam oder Auftraggeber die Möglichkeit bietet, den Test in Echtzeit zu verfolgen.

Der Testraum ist mit verschiedenen Kameras und Mikrophonen zur Aufzeichnung der Probandeninteraktion ausgestattet. Der Einsatz mehrerer Kameras resultiert aus der Tatsache, dass beim Benutzertest verschiedene Blickwinkel der Aufgabenbearbeitung von Interesse sind und aufgenommen werden sollten. So sind Aufnahmen der Arm- und Handbewegungen sowie des Gesichtsausdrucks essentielle Informationsquellen, um die Interaktion des Probanden vollständig zu erfassen /32/. Bei der Evaluation von Bildschirmsystemen ist das direkte Abgreifen und Aufzeichnen des Monitorsignals eine weitere wichtige Informationsquelle. Solche Daten können direkt über Protokollierungssoftware ausgewertet und visualisiert werden. In modernen Usability-Laboratorien werden auch Eye-Tracking Systeme zur Datenerhebung eingesetzt. Näheres zu Eye Tracking Systemen wird in Kapitel 4.2.9 aufgezeigt.

Im Kontrollraum oder auch Regieraum wird der Test durch den Testleiter überwacht und koordiniert. Er beinhaltet die zur Durchführung und Kontrolle notwendigen Monitore und Eingabesysteme sowie die Kamerasteuerung. Meist wird der Testleiter durch ein Team von Technikern bei der Durchführung des Tests unterstützt, so dass er sich primär auf die Beobachtung und ggf. Anleitung des Probanden konzentrieren kann.

In einer Analyse von 13 Usability-Labs aus dem Jahre 1994 wurden einige Eckdaten, wie die Anzahl an Kameras und die durchschnittliche Raumgröße verschiedener Usability Labs durch Nielsen /68/ ermittelt. Tabelle 4.1 zeigt die Ergebnisse dieser Studie. In fast allen untersuchten Usability-Labs wurden zum damaligen Zeitpunkt mindestens zwei Kameras verwendet. Das untermauert die These, dass eine vollständige Beobachtung der Probandeninteraktion nur durch den Einsatz mehrere Kameras möglich ist, die verschiedene Aufnahmepositionen und Blickwinkel während des Tests aufzeichnen können.

Company Name	Main Product	Other Labs in Company?	Date of First Usability Lab in Company	Floor Space of Typical Subject Room in Sq. Meters	Floor Space of Total Lab Area in Sq. Meters	Number of Rooms (Subject Rooms, Control Rooms, etc.)	Number of Cameras in Typical Subject Room	Scan Converter Used to Directly Tape Screen Image?	One-Way Mirror?	Usability Staff Supporting vs. Utilizing Lab
Ameritech	Communications service	No	1989	12.5	237	7	2	Yes	Yes	1 / 10
Bellcore	Telco software	Yes	1985	12.3	121	7	2	No	Yes	0.3 / 30
BT (British Telecom)	Telephone service	No	1988	40	96	3	3	Yes	Yes	0.5 / 70
IBM	Computer systems	Yes	1981	11.7	165	14	2	No	Yes	0.1 / 4
MAYA Design Group	Design consultants	No	1990	8.8	42.9	3	2	No	Yes	3 / 12
Microsoft Corp.	PC software	No	1989	10.8	181.3	19	2	Yes	Yes	4 / 22
NCR	Computer systems	Yes	1966	13.4	31.2	3	2	Yes	Yes	2 / 15
Philips, Corp. Design	Consumer electronics	Yes	1990	30	40	2	3	No	Yes	0.25 / 10
Philips, IPO	Consumer electronics	Yes	1990	9	35	3	2	No	No	1 / 25
SAP	Enterprise business apps	No	1992	37	63	2	1	No	Yes	2 / 12
SunSoft	Workstation software	No	1988	25.1	202.3	8	3	Yes	Yes	3.5 / 8
Symantec	PC software	No	1992	23.8	47.6	2	2	Yes	Yes	1 / 1
Taligent	PC operating systems	No	1992	13.4	26.8	2	2	No	Yes	1 / 2
Mean		38%	1987	19.1	99.2	5.8	2.2	46%	92%	1.5 / 17.0
Median		No	1989	13.4	63.8	3	2	No	Yes	1 / 12

Tabelle 4.1: Ergebnisse einer Untersuchung von 13 Usability-Labs /68/

Das Usability-Lab stellt, je nach Ausstattung, ein mächtiges Instrumentarium an möglichen Erhebungsmethoden innerhalb der Usability-Evaluation zur Verfügung. Datenhandschuhe ermöglichen die Messung von Puls und Hautleitwiderstand einer Testperson. Dadurch können schwer quantifizierbare Größen wie der Stress einer Versuchsperson gemessen werden. Sogenannte Performancedaten wie Eingabeprotokolle, Ausführungszeiten, Fehlerraten und Suchverhalten können über Eye-Tracking gemessen und über entsprechende Softwarepakete direkt ausgewertet werden.

Nielsen bemerkt zu den Usability-Labs entsprechend, dass solche hochtechnisierten Labore sicherlich eine komfortable Lösung zur Erhebung von Daten beim Benutzertest darstellen, jedoch keinesfalls eine zwingende Voraussetzung für einen Usability-Test sind.

"*I should stress, however, that special laboratories are a convenience but not an absolute necessity for usability testing. It is possible to convert a regular office temporarily into a usability laboratory, and it is possible to perform usability testing with no more equipment than a notepad.*" /21/

Usability-Labs sollten unter Berücksichtigung des Nutzungskontexts des zu bewertenden Produkts entsprechend angepasst und ausgestattet sein. Das bedeutet, dass die reale Arbeitssituation des späteren Nutzers im Usability-Lab so authentisch wie möglich abgebildet wird. Dies mag zwar bis zu einem gewissen Grade möglich sein, jedoch sind vollständig realistische Bedingungen in der Praxis nur schwer darzustellen. Umgebungseinflüsse, wie beispielweise die Unterbrechung der Arbeit durch Telefonanrufe

und sonstige Störungen wie Umgebungsgeräusche können in einem Laborsetting nicht oder nur teilweise nachgebildet werden.

4.2.3 Das mobile Usability-Lab

Für den Fall das kein Usability-Lab zur Verfügung steht oder die Evaluation in einem Usability-Lab nicht möglich ist, können mobile Lösungen zur Interaktionsaufzeichnung während der Evaluierung eingesetzt werden. Mit solch einem transportablen Usability-Lab ist es möglich, Benutzertests vor Ort, also am realen Arbeitsplatz oder Einsatzort, durchzuführen. Verschiedene Kameras und Mikrophone zur Interaktionsaufzeichnung, sowie drahtlose Empfänger für Video- und Audiosignale und Stative zur Arretierung des Equipments finden sich in dem in Bild 4.4 beschriebenen mobilen Usability-Lab.

Bild 4.4: Mobiles Usability Lab /54/

Der wesentliche Vorteil des mobilen Usability-Labs ist die Tatsache, dass der Test direkt vor Ort im realistischen Setting, der realen Arbeitsumgebung und somit dem wirklichen Nutzungskontext des Produkts, stattfindet. Die Laborsituation für den Probanden ist zwar weiterhin vorhanden, jedoch bei weitem nicht in dem Maße, wie im stationären Usability-Labor. In der Praxis sind mobile Usability-Labs das Mittel der Wahl, wenn Testpersonen außerhalb ihrer regulären Arbeit nicht zur Verfügung stehen oder der Testprototyp nicht transportabel ist. Schwere Industrieroboter und Werkzeugmaschinen, oder auch stationäre Kiosksysteme, wie Bankautomaten sind Beispiele nicht-transportabler Prototypen.

Eine Softwarelösung von Techsmith, die speziell für Usability-Tests entwickelt wurde erlaubt eine detaillierte Aufzeichnung, Observation und Analyse eines Usability-Tests. Mittels einer digitalen Kamera oder Webcam und eines handelsüblichen Computers, kann man bereits mit relativ geringem Investitionsaufwand schnell und einfach eine Evaluation überwachen und auswerten. Das Softwarepaket mit dem Namen Morae /41/ ist ab 1300€ erhältlich und, verglichen mit einem Usability-Test im Labor, eine kostengünstige Alternative zur Durchführung eines Usability-Tests.

4.2.4 Detaillierter Ablauf des Verfahrens

Bei einem Usability-Test wird zwischen einer Vorbereitungsphase und der Durchführungsphase unterschieden.

Die Vorbereitungsphase umfasst folgende Teilschritte:

- Festlegung der Fragestellung/Zielsetzung der Evaluation
- Auswahl der Testaufgaben
- Auswahl der Nutzer
- Technische und organisatorische Vorbereitung der Evaluation

Festlegung der Fragestellung/Zielsetzung der Evaluation

Zuerst sollte geklärt werden, welches Ziel der Usability-Test hat. Die Zielsetzung des Tests gibt wiederrum die Rahmenbedingungen und den Ablauf der Evaluation vor.

Mögliche Zielsetzung im Rahmen eines Usability-Tests sind beispielsweise:

- Abgleich und Überprüfung von Ergebnissen analytischer Evaluationen hinsichtlich ihrer Vorhersagekraft.
- Ermittlung von echten Problemen realer Benutzer.
- Ermittlung von Akzeptanzproblemen.
- Erhebung von weiteren Anforderungen, fehlenden Funktionalitäten und Verbesserungsvorschlägen.
- Vergleich verschiedener Produkte unter realistischen Einsatzbedingungen.

- Messung der Performance, also Erhebung quantitativer Daten.

Auswahl der Testaufgaben

Die Testaufgaben sollten möglichst realistisch sein und repräsentativ den späteren Arbeitsalltag der Nutzer widerspiegeln. Idealerweise bestehen die Arbeitsaufgaben aus echten Praxisaufgaben und fließen in voller Bandbreite in die Evaluation mit ein. Bei komplexen Systemen ist in der Regel keine Evaluation aller potentiellen Arbeitsaufgaben des Systems möglich, daher müssen selektiv Aufgaben ausgewählt werden, die repräsentativ und möglichst realistisch den späteren Nutzungskontext des Produkts darstellen.

Sollten unlösbare Probleme bei der Aufgabenbearbeitung auftreten, kann Hilfestellung durch den Testleiter geben werden /20/. Dies kann durch Hinweise auf ein Überdenken des aktuellen Lösungsweges oder Fragen nach dem aktuellen Benutzerziel durch den Testleiter geschehen. Wenn der Proband dennoch nach weiterem Suchen einer Lösung keinen Erfolg hat, wird die Aufgabenbearbeitung abgebrochen.

Auswahl der Testpersonen

Der Nutzungskontext, in dem das Produkt oder das technische System später eingesetzt werden soll, bestimmt die Auswahl der Testpersonen. Da zu jedem technischen Produkt auch eine entsprechende Zielgruppe existiert, sollte die Rekrutierung der Testpersonen eben aus dieser erfolgen. Handelt es sich bei dem technischen System um ein hochspezialisiertes 3D-Konstruktionsprogramm sollten entsprechende spätere Endnutzer, also zum Beispiel Konstruktionsingenieure, als Testpersonen ausgewählt werden. Handelt es sich um ein Profigerät im Handwerk, beispielsweise eine Bohrmaschine, sollte die Testgruppe aus Handwerkern bestehen, die im Alltag mit Produkten dieser Art arbeiten.

Technische und organisatorische Vorbereitung

Die technische und organisatorische Vorbereitung der Evaluation hängt in erster Linie von dem Evaluationsort ab. Es stellt sich hier die Frage, ob das technische System oder Produkt vor Ort, in seiner natürlichen Arbeitsumgebung getestet werden kann, oder ob die Evaluation im Usability-Lab stattfinden soll.

4.2.4.1 Optimale Anzahl an Probanden

Nielsen /42/ empfiehlt den Einsatz von fünf Probanden bei einem Usability-Test, da diese Zahl das beste Verhältnis von Investitionen zu Erkenntnisgewinn darstellt.

Das Ergebnis einer Formalisierung des Zusammenhangs von gefundenen Usability-Problemen und der Anzahl der eingesetzten Probanden /15/ bei einem Benutzertest zeigt Bild 4.5.

Ein erster Proband findet im Schnitt 35% der Gesamtzahl an Usability-Problemen in einem Interface. Danach sinkt der Erkenntnisgewinn mit steigender Anzahl der Probanden. Theoretisch sind für die vollständige Erfassung aller Probleme mindestens 15 Probanden von Nöten.

Im Zuge eines iterativen Gestaltungsprozesses, lautet Nielsens Empfehlung in diesem Zusammenhang, drei Testtermine mit jeweils 5 Probanden durchzuführen und nach dem Test die gefundenen Usability-Probleme zu beseitigen, anstatt einen Test mit 15 Probanden durchzuführen.

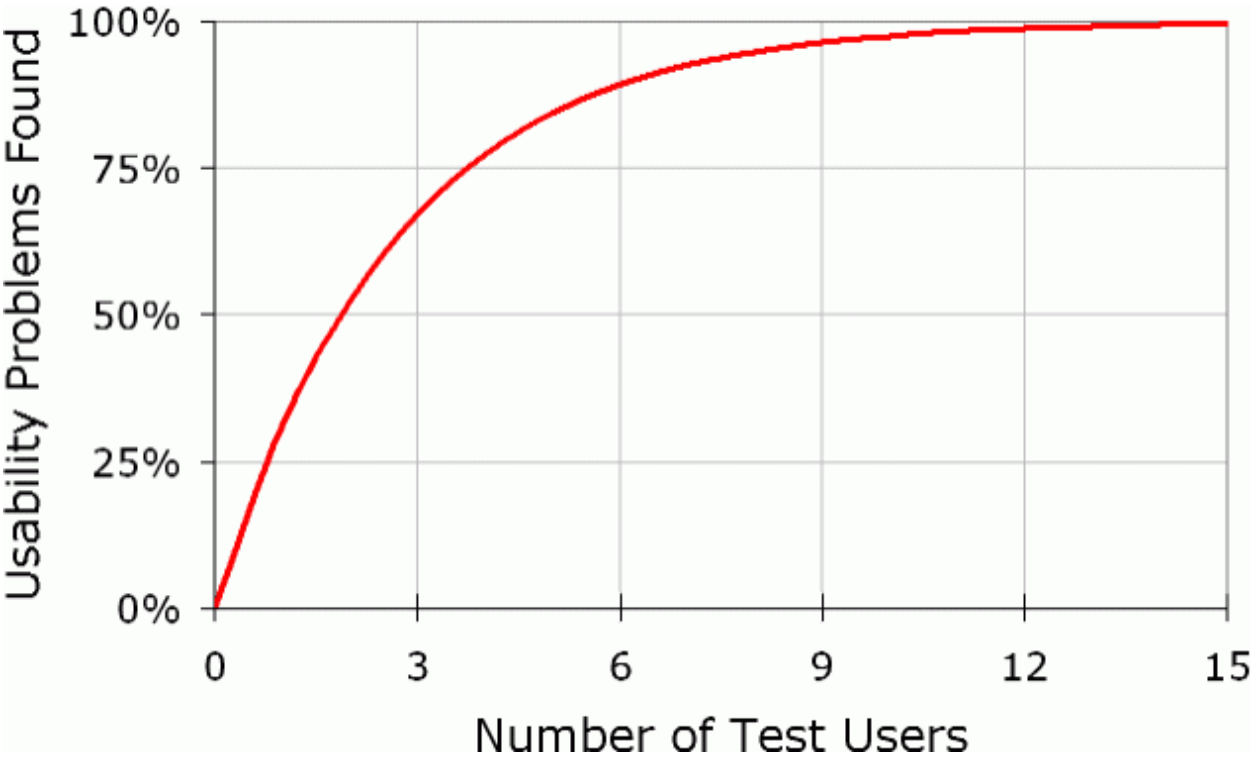

Bild 4.5: Problemfindungskurve – Benutzertest /55/

4.2.5 Mögliche Erhebungsmethoden beim Benutzertest

Es existiert heutzutage eine ganze Reihe von verschiedenen Erhebungsmethoden, die innerhalb eines Usability-Tests eingesetzt werden. Im folgenden Kapitel werden die wichtigsten und gängigsten Erhebungsmethoden vorgestellt und diskutiert. Es folgen Beispiele und eine Erläuterung der Vor- und Nachteile der jeweiligen Erhebungsmethode. Die Auswahl erfolgt nach /32/, /38/, /20/, /24/.

- Methode des lauten Denkens/Think-Aloud-Protocol

- Videoanalyse/ Videofeedback
- Eingabeprotokolle/Logfiles
- Aufmerksamkeitsanalyse/Eye-Tracking
- Interviews und Fragebögen
- Observation

Die verschiedenen Erhebungsmethoden werden in der Praxis oft miteinander kombiniert. Es macht auch durchaus Sinn einen Usability-Test, innerhalb dessen sich die Benutzer intensiv mit dem zu evaluierenden Produkt auseinandergesetzt haben, mit einen Interview oder Fragebogen zu kombinieren. Die oben aufgeführten Erhebungsmethoden werden im folgenden Kapitel separat beschrieben und tendenzielle Vor- und Nachteile aufgezeigt.

4.2.6 Methode des lauten Denkens/Think-Aloud-Protocol

Die Methode des lauten Denkens wurde traditionell als psychologische Forschungsmethode benutzt und ist mittlerweile eine verbreitete Methode zur Evaluation von Mensch-Maschine-Schnittstellen in der Praxis /32/, /20/.

4.2.6.1 Theorie und Beschreibung

“Lautes Denken“ ist eine Methode zur Erfassung bewusster handlungsbegleitender Kognitionen und Emotionen und bedeutet, dass der Proband während eines Benutzertests seine Gedanken und Kognitionen äußert /9/. Diese Art der Verbalisierung von Problemen der Testperson bei der Aufgabenbearbeitung ermöglicht eine unmittelbare Erhebung von subjektiven, qualitativen Daten durch den Versuchsleiter. Die Interpretation einer Mensch-Maschine-Schnittstelle durch einen Probanden, also das individuelle Verständnis einer Benutzeroberfläche durch einen späteren Nutzer, ist eine sehr effektive Methode /20/ um akute Usability-Probleme zu identifizieren. Die Methode des lauten Denkens, ist in diesem Sinne das empirische Pendant zum in Kapitel 4.4 beschriebenen Cognitive Walkthrough, der das Verhalten der Probanden und deren Problemlöseprozess durch Experten zu simulieren versucht.

Ebenso eignen sich die qualitativen Aussagen und Meinungen der Nutzer zur direkten, unverfälschten Verwendung innerhalb eines Evaluationsberichtes, da sie dessen Verständlichkeit und Erinnerungswert gegenüber dem jeweiligen Auftraggeber erheblich erhöhen können.

Die Methode des lauten Denkens bringt jedoch auch einige Nachteile mit sich. Als erstes wäre hier die Doppelbelastung des Probanden zu nennen. Er muss gleichzeitig die Testaufgaben lösen und “laut Denken“. Daraus resultiert eine langsamere Aufgabenbearbeitung durch die Probanden, was bedeutet, dass Performance Messungen bei Einsatz dieser Methode praktisch wertlos sind. Vorsichtig ist nach Nielsen mit Interpretatio-

nen der Probanden bei der Aufgabenbearbeitung umzugehen. Kausalitätszusammenhänge eines Problems und dessen Ursache durch die Nutzer sollten kritisch betrachtet und keinesfalls schlichtweg hingenommen werden. In solchen Situation, in denen Nutzer eine Theorie zur Ursache eines Problems formulieren, sollte der Testleiter sich, nach Nielsen, lieber auf den Part konzentrieren, in dem der Nutzer sich mit dem Problem beschäftigt hat.

Folgendes kurzes Beispiel /22/ soll das beschriebene Problem verdeutlichen:

Innerhalb eines Usability-Tests eines Benutzerinterfaces wurde zu Beginn des Testverlaufs durch den Beobachter festgestellt, dass ein spezifisches Dialogfeld durch den Probanden übersehen und nicht gefunden wurde. Im weiteren Testverlauf findet der Proband schließlich den gesuchten Dialog und mutmaßt direkt, dass er den gesuchten Dialog viel schneller gefunden hätte, wenn er an einer anderen Stelle im Benutzerinterface platziert worden wäre.

Die Vermutung des Probanden, eine andere Platzierung des Dialogfeldes im Interface wäre von Vorteil, mag zwar stimmen, jedoch sollte durch den Beobachter das Verhalten des Probanden dokumentiert werden, das in der Zeitspanne zwischen übersehenem Dialogfeld und letztendlichem Auffinden des Feldes zu beobachten war.

Der größte Vorteil der Methode des lauten Denkens ist das direkte Aufzeigen und Dokumentieren des Probandenverhaltens während der Aufgabenbearbeitung. Es kann nachvollzogen werden warum und wie sich die Probanden während eines Tests verhalten. Ein weiterer Vorteil der Methode ist neben der direkten, unverfälschten Dokumentation der Probandenaussagen der geringe technische und finanzielle Aufwand innerhalb der Anwendung. Da eine schriftliche Dokumentation der Probandenaussagen durch den Testleiter vollkommen ausreichend ist, werden keine technischen Hilfsmittel oder aufwändige Laborausstattung für den Test benötigt.

Die Tatsache, dass Probanden bei der Aufgabenbearbeitung Teilfunktionen oder Aspekte des Interfaces mit einer persönlichen Bewertung kommentieren, ist als Vorteil und Alleinstellungsmerkmal der Methode zu erwähnen. Zusätzlich werden die Gedanken direkt und verzögerungsfrei in der jeweiligen Aufgabensituation durch den Probanden verbalisiert. Eine spätere Rationalisierung der Äußerungen, wie zum Beispiel bei der Methode des Videofeedbacks (siehe Kapitel 4.2.7), wird dadurch ausgeschlossen.

4.2.6.2 Allgemeines zur Anwendung

Es ist im Allgemeinen ungewohnt und problematisch für die Probanden während der Aufgabenbearbeitung kontinuierlich ihre Gedanken zu verbalisieren. Gerade erfahrene Nutzer, die viele Interaktionsschritte selbstverständlich aus ihrer Erfahrung tätigen und

daher nicht kommentieren, müssen durch den Versuchsleiter immer wieder durch Rückfragen zum “Lauten Denken“ angewiesen werden.

Beispiele für Fragen durch den Testleiter sind /32/, /20/:

- Was denken Sie gerade?
- Was denken Sie, was diese Meldung zu bedeuten hat? (nachdem der Benutzer ein Feedback bekommen hat und augenscheinlich darüber nachdenkt was es zu bedeuten hat)
- Was denken Sie, was passieren wird?
- Welche Reaktionen haben Sie erwartet?

Es ist wichtig darauf zu achten, dass auf Fragen der Probanden zur Aufgabenbearbeitung mit einer Gegenfrage geantwortet wird:

- Frage des Probanden: Kann ich hier draufdrücken?
- Antwort des Testleiters: Was erwarten Sie, wenn Sie auf diesen Knopf drücken?

Unter Beachtung des Prinzips der selbständigen Aufgabenlösung durch den Probanden, sollten auch vorwegnehmenden Fragen wie

- Was denken Sie, was der Hinweis in der oberen rechten Ecke des Bildschirms zu bedeuten hat?

vermieden werden.

Um die Methode des lauten Denkens den Probanden anschaulich zu erklären und ihnen Ängste und Unsicherheit zu nehmen, schlägt Nielsen das Vorstellen eines kurzen Beispiels bei der Probandeneinweisung innerhalb der Testeinführung vor. Der Testleiter kann eine alltägliche Aufgabe, wie beispielweise das Nachschlagen eines Wortes im Duden, gegenüber den Probanden unter Verwendung der Methode des lauten Denkens vorführen.

4.2.7 Videoanalyse/Videofeedback

Die Methode der Videoanalyse ist, neben der Methode des lauten Denkens eine weitere, in der Praxis weit verbreitete Methode zur Erhebung und Dokumentation des Probandenverhaltens beim Usability-Test. Beide Methoden können selbstverständlich auch kombiniert werden. Nach Mackay /18/ ist der Einsatz von Videoaufzeichnungen bei Benutzertests ein unverzichtbarer Bestandteil zur detaillierten, minutiösen Erforschung der Mensch-Maschine-Kommunikation auf dem Gebiet der Interaktionsforschung.

4.2.7.1 Allgemeines

Unter der Erhebungsmethode der Videoanalyse versteht man die Dokumentation des Probandenverhaltens beim Usability-Test unter Einsatz einer oder mehrerer Kameras. In einer Studie /68/ zum Vergleich der Ausstattung von 13 industriell betriebenen Usability-Labs lag der Durchschnitt der eingesetzten Kameras pro Labor bei 2,2.

Werden zwei oder mehr Kameras eingesetzt, dienen diese meist dazu, Detailaufnahmen des Gesichts oder der Hände des Probanden aufzunehmen.

Die Protokollierung der Aufgabenbearbeitung der Probanden durch den Testleiter wird durch die Aufzeichnung auf Videoband erheblich erleichtert. <

4.2.7.2 Allgemeines zur Anwendung

Nielsen /23/ schätzt die Zeit für eine detaillierte Auswertung von Videoaufzeichnungen beim Usability-Test auf das Drei- bis Zehnfache der Zeit, die für die eigentliche Aufgabenbearbeitung durch den Probanden benötigt wird.

Das Videofeedback findet anschließend an den eigentlichen Usability-Test statt. Testleiter und Proband sichten gemeinsam das Videomaterial nach kritischen und unklaren Stellen innerhalb der Aufgabenbearbeitung. Das Ziel dieser Abschlusssitzung ist zum Beispiel die Klärung von Fragen, die der Testleiter während der Aufgabenbearbeitung notiert hat.

Die Methode der Videokonfrontation hat gegenüber der Methode des lauten Denkens den Vorteil, dass insbesondere bei einer komplexen Aufgabenbearbeitung, die Verbalisierungsleistung durch den Probanden entfällt und er seine volle Konzentration auf die Bearbeitung der Testaufgabe richten kann. Im Gegenzug birgt die Videoauswertung die Gefahr der Verhaltensrationalisierung /20/, /38/ durch die Tester, was die Reliabilität der Ergebnisse erheblich schmälern kann.

4.2.8 Eingabeprotokolle/Logfiles

Unter Logfile-Recording oder Eingabeprotokollen versteht man das automatische Computergesteuerte Mitschneiden von Probandeninteraktionen, sogenannter User Interface Events bei der Aufgabenbearbeitung in interaktiven Computersystemen. Die User Interface Events sind Systemereignisse, zum Beispiel Mausklicks, Mausbewegungen und Tastatureingaben die das Verhalten eines Nutzers auf der Benutzungsoberfläche einer Applikation widerspiegeln /11/.

Die Erhebung von Eingabeprotokollen durch automatisierte Software, die im Hintergrund während eines Benutzertests läuft, erlaubt eine objektive Informationssammlung

der Nutzerinteraktion innerhalb eines Bildschirmsystems. Diese automatisch generierten Eingabeprotokolle haben den entscheidenden Vorteil, dass menschliches Versagen bei der Datenerhebung praktisch ausgeschlossen werden kann. Eingabeprotokolle eignen sich hauptsächlich für die Evaluation von Webseiten und PC-Software und für den kompetitiven Vergleich von Gestaltungsalternativen. Sogenannte Vorher-Nachher-Vergleiche, etwa bei neugestalteten Dienstleistungsangeboten im Internet oder die Erhebung von Fehlerraten beim Ausfüllen elektronischer Bestellformulare sind typische Beispiele für die Verwendung von Eingabeprotokollen.

Nachteilig ist der vergleichsweise hohe technische Aufwand, der für die Vorbereitung und Nachbereitung der Datenerhebung zu betreiben ist. Darunter fallen beispielsweise die Erstellung und Programmierung einer entsprechenden Software zur Evaluation und die aufwändige Datenaufbereitung und Auswertung.

4.2.9 Aufmerksamkeitsanalyse/Eye-Tracking

Augen- und Blickbewegungen stellen die schnellsten Events in der Mensch-und Computerinteraktion dar /38/ und können daher nur mit speziellen Apparaturen, so genannten Eye-Trackern, zu Zwecken einer weiterführenden Analyse aufgezeichnet werden. Das Verwenden von Eye-Tracking bei einem Usability-Test ist aktuell das einzige Verfahren, mit dem der Blickverlauf eines Probanden beim Betrachten eines Interfaces registriert und aufgezeichnet werden kann.

Insbesondere im Bereich des Internets und der Webseitengestaltung, werden Eye-Trackingsysteme zur Evaluation eingesetzt. Der Grund hierfür ist, dass die Besucher einer Internetseite dieser sehr wenig Zeit geben, ihre Informationen zu vermitteln. Mittels Eye-Tracking können objektive Daten in Bezug auf das Wahrnehmungsverhalten eines Probanden beim Betrachten eines Interfaces gewonnen werden. So können beispielsweise Teilbereiche identifiziert werden, denen keine oder nur wenig Beachtung geschenkt wird.

Durch den Einsatz der Aufmerksamkeitsanalyse werden Bereiche identifiziert, die von den wesentlichen Aspekten eines Interfaces ablenken. Die Frage, auf welche Informationen sich ein Benutzer wie lange konzentriert und wie der Suchprozess eines Probanden auf einer Benutzeroberfläche abläuft, kann durch Eye-Tracking beantwortet werden.

Als größten Vorteil der Aufmerksamkeitsanalyse ist daher der erhobene Datentyp zu nennen. Durch den Einsatz der Aufmerksamkeitsanalyse können quantitative Daten in erheblichem Umfang gemessen werden und dies unter Ausschluss menschlicher Beobachtungsfehler. Nur unter Anwendung dieser Methode ist ein chronologisch und objektiv erhobener Blickbewegungspfad eines Probanden bei der Aufgabenbearbeitung in einem Interface messbar. Keine andere Methode kann dies leisten. Des Weiteren kön-

nen die erfassten Blickbewegungen durch entsprechende Software anschaulich visualisiert und auf das zu bewertende Interface projiziert werden. Durch diese Visualisierung der Blickbewegungs- und Fixaktionspfade, können Rückschlüsse auf die örtlich-räumliche Gestaltung (Design und Struktur der Informationselemente) der Mensch-Maschine-Schnittstelle gezogen werden.

Nachteilig ist das Fehlen von Standards in der Interpretation des Datenmaterials. Durch Eye-Tracking ist zwar feststellbar, welchen Interfacebereich ein Proband wie lange fixiert hat, allerdings nicht, warum der Proband diesen Bereich fixiert hat. Zudem müssen die vergleichsweise hohen Anschaffungskosten für Geräte und Software berücksichtigt werden. Schlussendlich gibt es aktuell noch Probleme bei der Anwendung des Verfahrens mit Brillenträgern.

4.2.9.1 Beispiele für Eye-Tracking

Bild 4.6 zeigt einen modernen Eyetracker der Firma Tobii. Der Anschaffungspreis eines solchen Gerätes inklusive der entsprechenden Auswertungssoftware beläuft sich auf ca. 50.000 US-Dollar /43/. Bild 4.7 zeigt ein Beispiel für die Auswertung einer Eye-Tracking Studie. Die Hotspots, also die Bereiche, die während der Interaktion oft und lange fixiert wurden, werden entsprechend farblich hervorgehoben.

Bild 4.6: Ein Eyetracker der Firma Tobii /56/

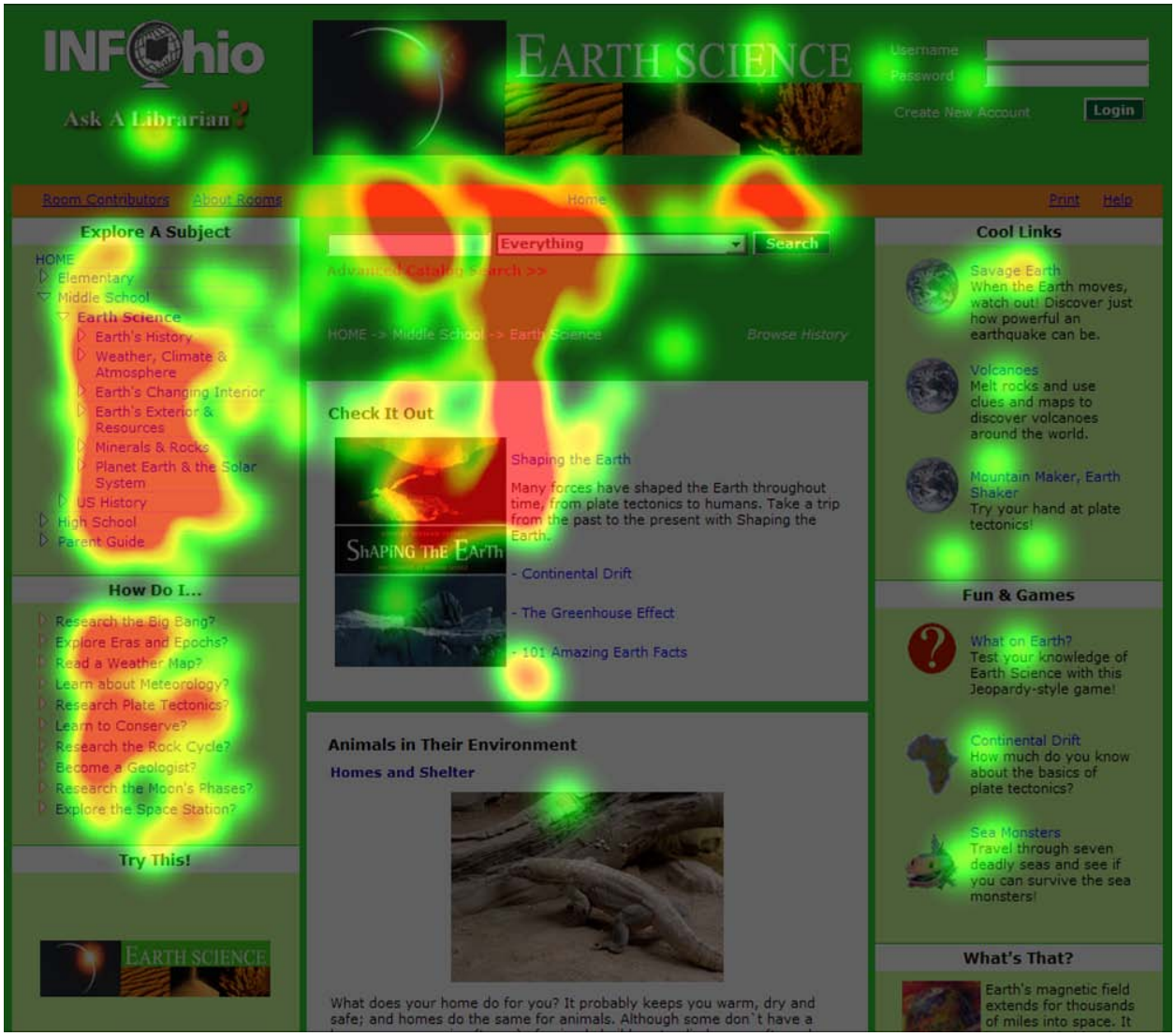
INFOhio
Ask A Librarian?
EARTH SCIENCE
Username
Password
Create New Account
Login
Room Contributors
About Rooms
Home
Print
Help
Explore A Subject
HOME
Elementary
Middle School
Earth Science
Earth's History
Weather, Climate & Atmosphere
Earth's Changing Interior
Earth's Exterior & Resources
Minerals & Rocks
Planet Earth & the Solar System
US History
High School
Parent Guide
Everything
Search
Advanced Catalog Search >>
HOME -> Middle School -> Earth Science
Browse History
Check It Out
Shaping the Earth
Many forces have shaped the Earth throughout time, from plate tectonics to humans. Take a trip from the past to the present with Shaping the Earth.
SHAPING THE EARTH
- Continental Drift
- The Greenhouse Effect
- 101 Amazing Earth Facts
Cool Links
Savage Earth
When the Earth moves, watch out! Discover just how powerful an earthquake can be.
Volcanoes
Melt rocks and use clues and maps to discover volcanoes around the world.
Mountain Maker, Earth Shaker
Try your hand at plate tectonics!
How Do I...
Research the Big Bang?
Explore Eras and Epochs?
Read a Weather Map?
Learn about Meteorology?
Research Plate Tectonics?
Learn to Conserve?
Research the Rock Cycle?
Become a Geologist?
Research the Moon's Phases?
Explore the Space Station?
Fun & Games
What on Earth?
Test your knowledge of Earth Science with this Jeopardy-style game!
Continental Drift
How much do you know about the basics of plate tectonics?
Sea Monsters
Travel through seven deadly seas and see if you can survive the sea monsters!
Try This!
EARTH SCIENCE
Animals in Their Environment
Homes and Shelter
What does your home do for you? It probably keeps you warm, dry and safe; and homes do the same for animals. Although some don't have a
What's That?
Earth's magnetic field extends for thousands of miles into space. It

Bild 4.7: Hotspots einer Eyetrackingstudie /57/

4.2.10 Fragebögen und standardisierte Interviews

Der Einsatz von Interviews und Fragebögen vor, während oder nach einem Benutzer-Test ist eine weit verbreitete Erhebungsmethode, um bereits erhobene Daten aus dem Test durch subjektive Aussagen der Probanden zu ergänzen.

4.2.10.1 Allgemeines zum Fragebogenverfahren

Die Elemente eines Fragebogens nennt man Items. Als Items bezeichnet man Fragen, Behauptungen oder Aufgaben /14/. Items gleicher Themengruppen werden innerhalb eines Fragebogens in sogenannte Subskalen aufgeteilt.

Die Beantwortungsmöglichkeiten des Probanden hängen vom verwendeten Fragebogentypen ab. Häufig verwendete Beantwortungsmöglichkeiten sind semantische Diffe-

rentiale, also gegensätzliche Wortpaare oder abgestufte Ratingskalen. Unter einem semantischen Differential versteht man zwei gegensätzliche Adjektive, wie beispielsweise praktisch – unpraktisch oder menschlich - technisch.

Im AttrakDiff /44/, einem Fragebogen zur Messung der “wahrgenommenen hedonischen und pragmatischen Qualität“ von interaktiven Produkten, werden solche semantische Differentiale innerhalb einer siebenstufigen Ratingskala verwendet.

ISONORM 9241/10 /45/, ein Fragebogen zur Beurteilung von Software auf Grundlage der Ergonomie-Norm ISO 9241 Teil:10 verwendet ebenfalls eine siebenstufige bipolare Ratingskala unter Verwendung von semantischen Differentialen.

Neben diesen abgestuften, bipolaren Skalen können auch vorgegebene Antwortoptionen in Form von Multiple-Choice-Antworten oder freien Texten gegeben werden.

4.2.10.2 Fragebögen – Vor- und Nachteile

Die größte Stärke des Fragebogenverfahrens ist sicherlich, dass mit relativ geringen materiellen Aufwand, eine große Anzahl an subjektiven Benutzermeinungen standardisiert abgefragt und einfach ausgewertet werden kann.

Die Voraussetzung zur Anwendung des Fragebogenverfahrens ist demnach, dass die Probanden hinreichende Erfahrung mit dem technischen Produkt oder interaktiven System haben. Die Auswertung von vollstandardisierten Fragebögen, wie z.B. des ISONORM 9241/10, ist unproblematisch und schnell zu realisieren. Normalerweise wird über Mittelwertbildung und Standardabweichung ein graphisches Profil der Stärken und Schwächen eines Systems erstellt.

Sardonick /34/ weist auf folgende Chancen und Risiken von Fragebögen als Evaluationsmethode beim Usability-Test hin:

Chancen

- Kaum Beeinflussung durch Dritte (z.B. Interviewer).
- Klare Trennung von Evaluationsaspekten durch Strukturierung der Fragebögen möglich.
- Subjektive Daten können kategorisiert und statistisch analysiert werden.
- Fragebogen können relativ unaufwändig an jeder Stelle des Entwicklungsprozesses eingesetzt werden.
- Es liegt eine Vielzahl an standardisierten Fragebögen zur Usability-Evaluation vor.
- Möglichkeit einer großen Stichprobe durch Onlineerhebungen.

Risiken

- Je strukturierter das Verfahren, desto weniger Freiraum bleibt für den Nutzer.
- Fragen fokussieren auf zuvor ausgewählte Bereiche, daher Einschränkung der Bewertungsfreiheit.
- Qualität der Daten abhängig von der Qualität des Erhebungsinstruments
- Die Entwicklung von neuen Fragebögen nach den Regeln der wissenschaftlichen Testkonstruktion(in Bezug auf Validität, Reliabilität und Objektivität) ist sehr aufwendig.

Im Folgenden werden drei standardisierte Fragebogen zur Messung verschiedener Software-ergonomischer Aspekte vorgestellt. Hierbei wurden

4.2.10.3 Attrakdiff – Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität

AttrakDiff ist ein Instrument zur Messung der Attraktivität interaktiver Produkte. Die Bewertung erfolgt über ein siebenstufiges, semantisches Differential gegensätzlicher Adjektivpaaren. Der Fragebogen besteht aktuell aus insgesamt 28 Items /44/.

Der Fragebogen AttrakDiff führt neue Qualitätskriterien für die Bewertung von interaktiven Produkten ein. Nach Hassenzahl /10/ widmet sich das bisher veröffentlichte Instrumentarium der veröffentlichen Fragebögen zu stark den Qualitätskriterien Effektivität und Effizienz der ISO 9241-11. Die Messung der Zufriedenstellung, wird weitgehend außer Acht gelassen und lässt sich letztlich auch meistens auf Maße der Effizienz und Effektivität zurückführen.

Innerhalb des zugrunde liegenden Modells des Fragebogens werden drei Nutzerqualitäten definiert und separat betrachtet:

- Pragmatische Qualität (z.B. praktisch, handhabbar)
- Hedonische Qualität – Bereich Stimulation (kreativ, herausfordernd)
- Hedonische Qualität - Bereich Identität (fachmännisch, verbindend)

Innerhalb des zugrundeliegenden, theoretischen Arbeitsmodells werden vier Aspekte definiert:

- Durch den Gestalter intendierte Produktqualität
- Subjektive Qualitätswahrnehmung und -bewertung
- Voneinander unabhängige pragmatische und hedonische Qualität
- Verhaltens- und emotionale Konsequenzen

Das Arbeitsmodell des AttrakDiff ist in Bild 4.8 dargestellt.

Die pragmatische Qualität bedeutet, die Fähigkeit eines interaktiven Produktes, das Bedürfnis des Nutzers nach Zielerreichung, durch die Bereitstellung nützlicher und be-

nutzbarer Funktionen zu befriedigen. Die hedonische Qualität ist in die zwei Gruppen Stimulation und Identität aufgeteilt..

Die Wahrnehmung der verschiedenen Qualitäten führt letztlich zur Gesamtbewertung der Attraktivität des untersuchten Produkts:

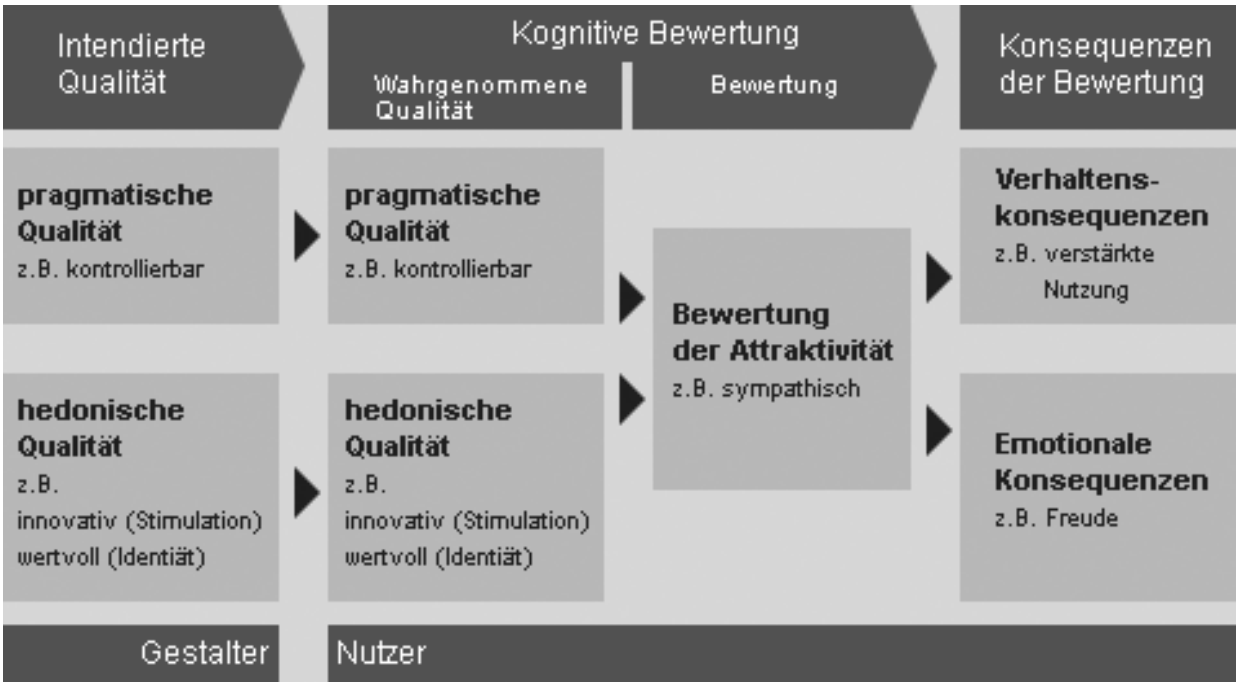

Bild 4.8: Arbeitsmodell des AttrakDiff /58/

AttrakDiff ist ein interessanter Fragebogen, da das ihm zugrunde liegenden Modell einen Versuch darstellt, die Messbarkeit der Qualität eines Produktes über die Gebrauchstauglichkeit hinaus mit neuen Kriterien zu erweitern.

Der Fragebogen ist kostenlos im Internet verfügbar /44/ und es wird zusätzlich die Möglichkeit der automatisierten Auswertung angeboten.

Die Ergebnisse der Probandenbefragung werden in ein Portfolio mit Konfidenzintervallen eingeordnet, wodurch die Positionierung des Produkts innerhalb der definierten Qualitäten erkennbar ist. Zusätzlich werden die Mittelwerte der semantischen Differentiale ausgewertet und ein Bewertungsprofil der Verteilung auf die einzelnen Wortpaare gebildet. Darüber hinaus werden Erläuterungen und Interpretationshilfen zu den Ergebnissen des Fragebogens gegeben. Es besteht auch die Möglichkeit, verschiedene Produkte innerhalb derselben Bewertung darzustellen, siehe Bild 4.9, Bild 4.10 und Bild 4.11.

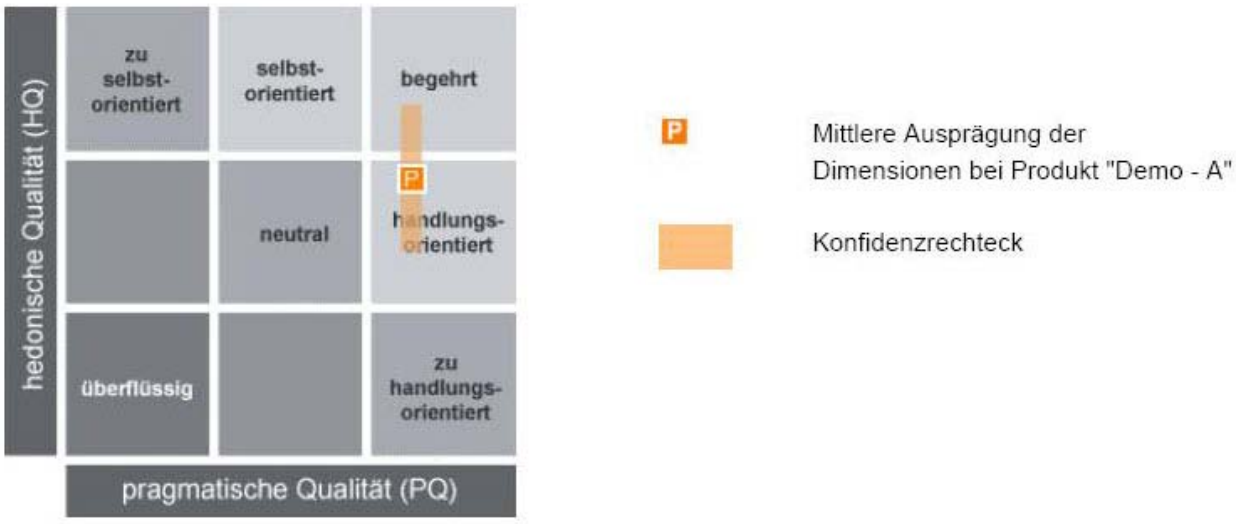

Bild 4.9: Portfolio der “Charakterbereiche“ des untersuchten Produkts /59/

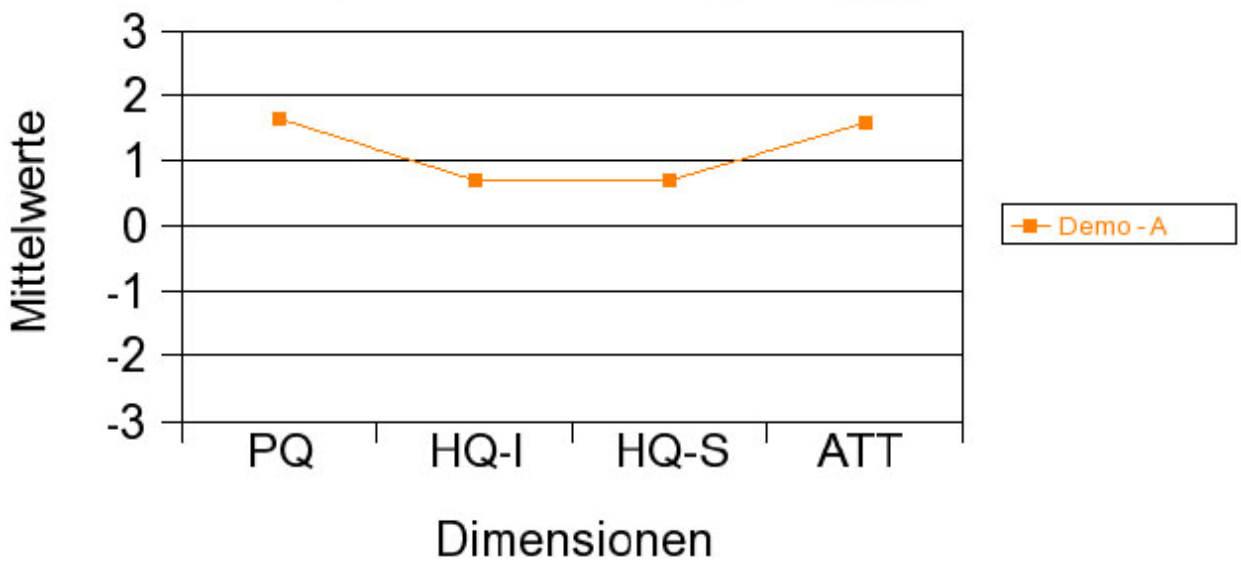

Bild 4.10: Diagramm der Mittelwerte /60/

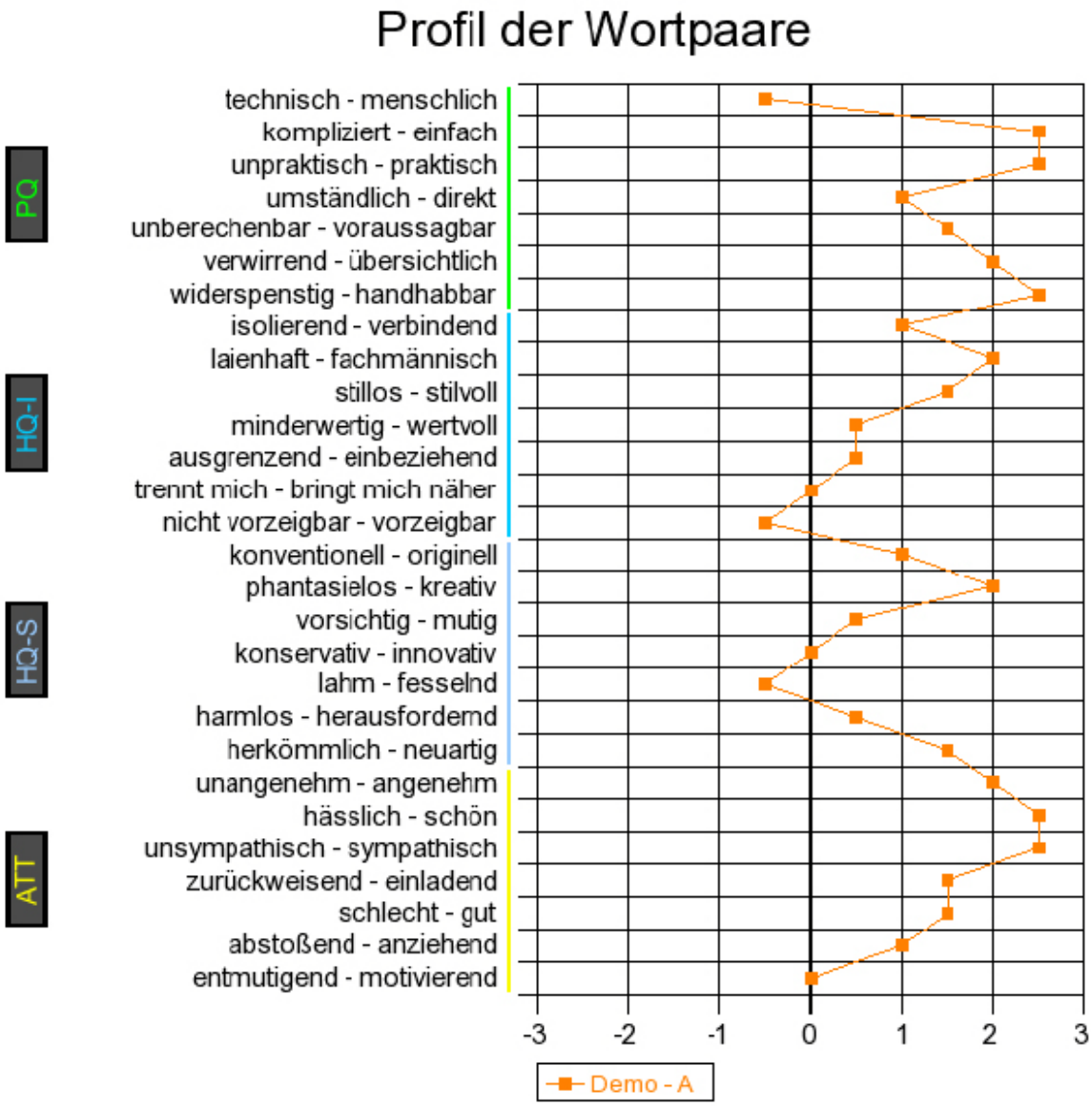

Bild 4.11: Verteilung der Mittelwerte auf die semantischen Differentiale /61/

4.2.10.4 QUIS

Als weiteren Vertreter der Fragebogenverfahren stellt Shneiderman /39/ mit QUIS, dem Questionnaire for User Interface Satisfaction, einen Fragebogen zur Messung und Bewertung der subjektiven Zufriedenheit der Nutzer bei der Interaktion mit der Mensch-Maschine-Schnittstelle vor. Der Fragebogen existiert in einer Kurz- und Langversion und muss lizensiert werden.

Der Fragebogen wird von einer interdisziplinären Forschergruppe der Universität Maryland ständig weiterentwickelt. Die aktuellste Version des Fragebogens /46/ hat die Versionsnummer 7.0 und ist in fünf Sprachen, darunter auch Deutsch, erhältlich.

Der Fragebogen enthält einen demographischen Fragenteil und verschiedene Subskalen mit systemspezifischen Items, wie z.B. Feedback der Software, Terminologie und Lernförderlichkeit. Durch den Probanden wird anhand eines semantischen Differentials innerhalb einer neunstufigen Skala die Ausprägung des betrachteten Items bewertet. Als Option gibt es noch die Möglichkeit innerhalb jedes Items nicht bewertbar anzukreuzen. Der Fragebogen ist so gestaltet, dass ganze Subskalen, jedoch nicht einzelne Items weggelassen werden können, z.B. für den Fall, dass noch kein Onlinehandbuch vorliegt. Der Proband kann zusätzlich, nach Bearbeitung jeder Subskala, freie Kommentare hinzufügen.

4.2.10.5 ISONORM 9241/10

Basierend auf der Vorgängerversion der bereits vorgestellten Norm 9241-110, existiert mit dem ISONORM 9241/10 /30/ ein Fragebogen zur Beurteilung von Software auf der Grundlage der Ergonomie-Norm ISO 9241 Teil: 10.

Als Subskalen dienen dem ISONORM 9241/10 die sieben Gestaltungsanforderungen der ISO Norm 9241-10. Jede Subskala umfasst fünf Items, die anhand einer siebenstufigen Skala bewertet werden. Zusätzlich werden noch einige soziodemographische Merkmale wie Alter, Geschlecht, Beruf und Kenntnisse am Computer abgefragt.

Die Auswertung des Fragebogens erfolgt über Mittelwertbildung aller Items innerhalb einer Subskala. Dadurch lassen sich Hinweise gewinnen, in welchen Bereichen der sieben Gestaltungsanforderungen Verbesserungspotential liegt. Wie alle standardisierten Fragebögen, sind auch die Items des ISONORM 9241/10 sehr allgemein formuliert und geben keine Hinweise auf konkrete Usability-Probleme oder Lösungsvorschläge. Bild 4.12 zeigt exemplarisch die Subskala Steuerbarkeit des Fragebogens.

Steuerbarkeit

Können Sie als Benutzer die Art und Weise, wie Sie mit der Software arbeiten, beeinflussen?

Die Software ...	---	--	-	-/+	+	++	+++	*Die Software ...*
bietet keine Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.	☐	☐	☐	☐	☐	☐	☐	bietet die Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.
erzwingt eine unnötig starre Einhaltung von Bearbeitungsschritten.	☐	☐	☐	☐	☐	☐	☐	erzwingt keine unnötig starre Einhaltung von Bearbeitungsschritten.
ermöglicht keinen leichten Wechsel zwischen einzelnen Menüs oder Masken.	☐	☐	☐	☐	☐	☐	☐	ermöglicht einen leichten Wechsel zwischen einzelnen Menüs oder Masken.
ist so gestaltet, daß der Benutzer nicht beeinflussen kann, wie und welche Informationen am Bildschirm dargeboten werden.	☐	☐	☐	☐	☐	☐	☐	ist so gestaltet, daß der Benutzer beeinflussen kann, wie und welche Informationen am Bildschirm dargeboten werden.
erzwingt unnötige Unterbrechungen der Arbeit.	☐	☐	☐	☐	☐	☐	☐	erzwingt keine unnötigen Unterbrechungen der Arbeit.

ISONORM 9241/10

Bild 4.12: Fragebogen ISONORM 9241/10 – Subskala Steuerbarkeit /62/

4.2.10.6 Interviews

Interviews sind im Vergleich zu den Fragebogenverfahren deutlich flexibler in ihrer Anwendung, jedoch aufwändiger durchzuführen und auszuwerten, da im Allgemeinen keine standardisierten Antworten der Probanden zu erwarten sind. Normalerweise wird in einem Interview, dem Probanden nach der Aufgabebearbeitung eines vorangegangenen Usability-Tests, ein Satz von vorformulierten Fragen gestellt. Die Antworten werden durch den Interviewer dokumentiert.

Die Möglichkeit des Interviewers, während des Interviews auf Antworten des Probanden zu reagieren und, darauf Aufbauend, die Befragung zu modifizieren ist ein wesentlicher Vorteil dieser Methode. Gleichzeitig wird aber durch diese Möglichkeit der Variation innerhalb dieser Erhebungsmethode der Auswertungsaufwand erheblich gesteigert.

4.2.11 Observation

Bei der Observation wird das Verhalten des Probanden durch den Testleiter beobachtet. Die Observation der Probanden bei der Aufgabenbearbeitung durch den Testleiter ist sicher die kostengünstigste Erhebungsmethode beim Benutzertest. Der Testleiter nimmt hierbei eine passive Rolle ein, d.h. er hilft dem Probanden bei auftretenden Problemen nicht weiter und gibt keine Hilfestellung.

Treten während der Aufgabenbearbeitung kritische Ereignisse auf, so werden diese dokumentiert. Dabei kann es sich um Überlegen und Stocken während der Aufgabenbearbeitung oder um verbale Äußerungen des Probanden handeln.

Erfahrung und ein geschultes Auge des Testleiters sind für die Observation eines Probanden beim Usability-Test unerlässlich. Der Testleiter muss die Fähigkeit besitzen, das Verhalten des Probanden möglichst neutral und interpretationsfrei zu erfassen. Die Qualifikation des Testleiters entscheidet daher maßgeblich, über Qualität und das Endergebnis des Usability-Tests.

4.3 Heuristische Evaluation

Bei der heuristischen Evaluation untersuchen Usability Experten unter freiem, explorativem, voneinander unabhängigen Vorgehen eine Benutzerschnittstelle oder ein Produkt anhand von festgelegten Gestaltungsrichtlinien und Usability-Prinzipien, sogenannten Heuristiken.

Kategorisch zählt die Heuristische Evaluation zu den formal-analytischen und nicht zu den empirischen Verfahren.

Das Ziel einer heuristischen Evaluation ist es, möglichst alle Usability-Probleme des untersuchten Systems aufzudecken und diese zu kategorisieren. Dies erfolgt entweder funktional oder/und nach den gegebenen Heuristiken.

Durch die abschließende Analyse der quantitativen Verteilung der einzelnen Usability-Probleme mit einer individuellen Gewichtung, lassen sich Aussagen und Handlungsbedarf für die Umgestaltung des Systems ableiten.

4.3.1 Heuristiken

Die zehn Heuristiken wurden erstmals von Nielsen und Molich 1990 formuliert und von Nielsen 1994 erneut überarbeitet /26/ und drücken die erwünschte Eigenschaft der Interaktion zwischen dem Nutzer und einem interaktivem System aus /16/:

1. **Visibility of system status**

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world**

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom**

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards**

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. **Error prevention**

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. **Recognition rather than recall**

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use**

Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design**

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. **Help users recognize, diagnose, and recover from errors**

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation**

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Die deutsche Übersetzung nach Klemmert /13/:

1. **Sichtbarkeit des Systemstatus**

Das System sollte den Nutzer immer informieren, was gerade vorgeht, durch geeignetes Feedback innerhalb angemessener Zeit.

2. **Übereinstimmung zwischen System und realer Welt**

Das System soll die Sprache des Nutzers sprechen, mit Wörtern, Formulierungen und Konzepten, die dem Nutzer vertraut sind und nicht mit systemorientierten Begriffen. Folge Konventionen aus der Realität, sodass Informationen in natürlicher Reihenfolge erscheinen.

3. **Benutzerkontrolle und Freiheit**

Nutzer wählen Softwarefunktionen oft versehentlich. Sie brauchen einen klar markierten Notausgang, um einen unabsichtlich erreichten Zustand verlassen zu können, ohne durch einen ausgedehnten Dialog hindurch zu müssen. Unterstütze Undo und Redo.

4. **Konsistenz und Standards**

Die Nutzer sollten nicht überlegen müssen, ob verschiedene Begriffe, Situationen oder Aktionen dasselbe bedeuten. Folge Plattform-Konventionen.

5. **Fehler vermeiden**

Besser noch als gute Fehlermeldungen ist ein sorgfältiges Design, das verhindert, dass überhaupt erst Probleme auftreten.

6. **Erkennen vor Erinnern**

Mache Objekte, Aktionen und Optionen sichtbar. Der Nutzer sollte nicht gezwungen werden, sich Informationen aus einem Teil eines Dialogs für einen anderen Teil zu merken. Anleitungen zur Benutzung eines Systems sollten sichtbar oder leicht auffindbar sein.

7. **Flexibilität und effiziente Nutzung**

Beschleunigungsmöglichkeiten, die der ungeübte Nutzer nicht sieht, können für erfahrene Nutzer die Interaktionsgeschwindigkeit erhöhen, sodass das System sowohl für Neulinge als auch für Experten geeignet ist. Erlaube Nutzern, häufige Aktionen auf ihre Bedürfnisse zuzuschneiden.

8. **Ästhetisches und minimalistisches Design**

Dialoge sollten keine Information enthalten, die irrelevant ist oder selten benötigt wird. Jedes Extra an Information in einem Dialog konkurriert mit relevanten Informationen und vermindert deren relative Sichtbarkeit.

9. **Unterstützung beim Erkennen, Verstehen und Bearbeiten von Fehlern**

Fehlermeldungen sollten in klarer Sprache (keine Kodierungen) gegeben werden. Sie sollten das Problem genau beschreiben und konstruktiv eine Lösung vorschlagen.

10. Hilfe und Dokumentation

Obwohl es besser ist, wenn das System ohne Dokumentation benutzt werden kann, kann es nötig sein, Hilfe und Information mitzugeben. Jede solche Information sollte leicht zu durchsuchen zu sein, die Aufgaben des Nutzers in den Mittelpunkt stellen und konkrete Schritte zur Ausführung nennen. Die Dokumentation sollte nicht zu umfangreich sein.

4.3.2 Detaillierter Ablauf des Verfahrens

Nach Nielsen /24/ besteht die Heuristische Evaluation aus 4 Phasen:

- Vorevaluatives Training
- Evaluationsdurchgänge
- Auswertung der Usability-Probleme
- Bewertung und Kategorisierung aller gefundenen Hinweise

Die Zusammensetzung des Teams besteht aus einem Versuchsleiter und den Usability-Experten, deren Anzahl variabel ist. Näheres zur optimalen Anzahl an Evaluatoren bei der Heuristischen Evaluation findet sich in Kapitel 4.3.3.

1. Vorevaluatives Training

Diese Phase bereitet die eigentliche Evaluation vor. Der Versuchsleiter gibt den Experten einen Überblick über die zu verwendende Methode und Heuristiken, um eine einheitliche Terminologie sicherzustellen. Danach folgt ein kurzer Überblick zum Nutzungskontext und der Wissensdomäne des zu bewertenden Systems.

Die Experten werden anhand eines Szenarios, das den späteren Arbeitsalltag des tatsächlichen Endnutzers repräsentiert, oder das beispielhafte Vorlegen von Arbeitsaufgaben für die spätere Zielgruppe sensibilisiert, beispielsweise durch Erläuterung des grundsätzlichen Arbeitsablaufs und Bedienkonzepts.

2. Evaluationsdurchgänge

Mit Hilfe der Heuristiken gehen die Experten unabhängig voneinander das Arbeitsszenario oder die Aufgabenstellung im System durch und stellen Verstöße gegen selbige fest. Es sollten mindestens zwei Evaluationsdurchgänge durchgeführt werden /24/, da der erste Durchgang zur Gewöhnung an das Szenario und System selbst benötigt wird, der zweite eine möglichst vollständige und analytische Erfassung einzelner Probleme und Ursachen unterstützt.

Das bloße Feststellen und Auffinden von Usability-Problemen steht in dieser Phase im Vordergrund, und es sollen keine Lösungen erdacht werden. Wie die Experten durch

das System gehen ist ihnen überlassen, es existieren hierbei individuelle Freiheiten in Bezug auf die Reihenfolge der Aufgabenbearbeitung. Die Probleme werden auf einem Berichtsbogen notiert, der folgende Struktur aufweisen sollte: Benennung des Problems, Fundort im System, Beschreibung des Problems, Erwartete Auswirkung.

3. **Auswertung der Usability-Probleme**

Die Auswertung der Berichtsbögen der Experten obliegt dem Versuchsleiter. Seine Aufgabe ist die Erstellung einer überschneidungsfreien Gesamtliste der identifizierten Usability-Probleme aus den einzelnen Berichtsbögen der Experten. Hierbei können die Erfahrung und das Fingerspitzengefühl des Versuchsleiters maßgeblich über die Güte der Gesamtliste entscheiden /32/.

Der Versuchsleiter entscheidet auch über das Niveau der Granulation der Gesamtproblemliste, also den Detaillierungsgrad der Formulierung der identifizierten Probleme. Der Vorteil einer hohen Granulation ist die detaillierte Beschreibung jedes identifizierten Problems. Bei komplexen Evaluationen kann ein zu hohes Granulationsniveau jedoch zu unübersichtlichen Problemlisten führen. Daher sollte ein für das jeweilige Untersuchungsobjekt angemessenes Niveau der Granulation gewählt werden.

4. **Bewertung und Kategorisierung aller gefundenen Hinweise**

Das Ziel der Abschlussphase der Heuristischen Evaluation ist die Bewertung der Schwere der einzelnen Probleme der Gesamtliste mit Hilfe der Experten. Aus dieser priorisierten Gesamtproblemliste können nun letztlich Handlungsbedarf und Optimierungshinweise zur Verbesserung der Usability des untersuchten Produkts gewonnen werden. Es sollten alle Experten an der Bewertung teilnehmen, da dies die Qualität der Bewertung erhöht. Um Schwere und Fatalität der festgestellten Usability-Probleme der Gesamtliste zu identifizieren wird folgendes Schema nach Nielsen /24/ angewendet:

Die Schwere eines Usability-Problems ist eine Kombination aus den 3 Faktoren

- Frequenz des Auftretens
- Einfluss auf die Arbeitsabläufe
- Persistenz des Auftretens

Die individuellen Auswirkungen eines Usability-Problems auf Marktakzeptanz und Kundenwünsche sowie Unternehmensprestige und Eindruck werden in dem Zusatzfaktor

- Markteinfluss

berücksichtigt.

Die Gewichtung der Schwere der gefundenen Usability-Probleme erfolgt mit einer fünfstufigen Skala von null bis vier und wird in Tabelle 4.2 erläutert.

Schwere	Erläuterung
0	Ich stimme nicht zu, dass dies überhaupt ein Usability-Problem ist.
1	Kosmetisches Problem – Braucht nicht behoben zu werden, außer der Projektrahmen stellt genügende Zeit zur Verfügung.
2	Geringfügiges Usability-Problem – Der Behebung sollte untergeordnete Priorität gegeben werden.
3	Bedeutendes Usability-Problem – Es ist wichtig, es zu beheben und sollte daher eine hohe Priorität erhalten.
4	Usability Katastrophe – Problem muss behoben werden, bevor es ausgeliefert werden kann.

Tabelle 4.2: Schwereskala der Heuristischen Evaluation

Die Bewertung der Probleme erfolgt anhand der eingeführten Skala von Nielsen und ergibt die priorisierte Problemliste, die bereits erste Entscheidungshilfen und Optimierungspotentiale aufzeigen kann.

Bei einem technischen System oder Produkt bietet sich eine Kategorisierung in Subsysteme oder Baugruppen an. Aus der Verteilung der gewichteten Probleme in den Subkategorien lassen sich so leicht Schwachstellen des Systems identifizieren. In einer optionalen Debriefingsitzung können die Evaluationsergebnisse im Beisein von Experten, Systementwicklern und Versuchsleiter präsentiert und diskutiert werden um konkrete Lösungsvorschläge zu erarbeiten.

4.3.3 Optimale Anzahl an Experten

In einer Fallstudie von Nielsen /27/, in der 19 Experten, insgesamt 16 Usability-Probleme eines Sprachcomputers für Telefonbanking identifizieren, ergab sich folgende Verteilung in Bezug auf identifizierte Probleme pro Evaluator:

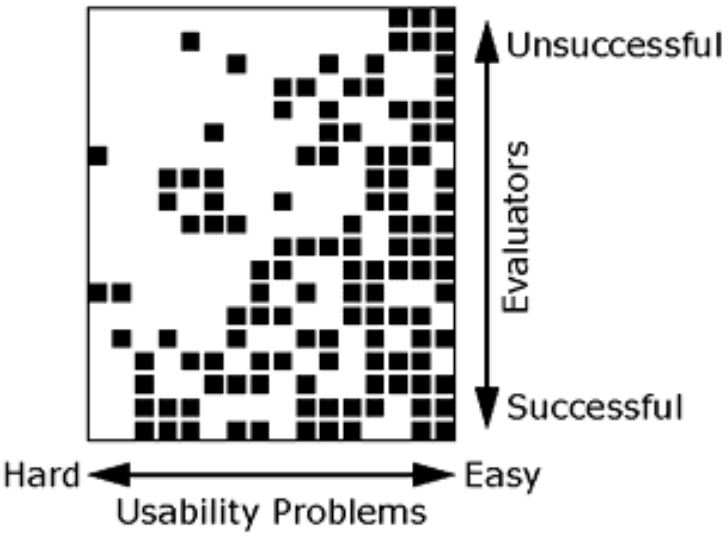

Bild 4.13: Verteilung: 19 Evaluatoren; 16 Usability Probleme /63/

Offensichtlich ist es keinem Experten gelungen alle Probleme zu finden, der “erfolgreichste“ konnte 10 von 16 Problemen identifizieren.

Bestimmte Usability-Probleme sind leichter zu finden, andere wiederum schwerer. Die Tatsache, dass manche Evaluatoren insgesamt zwar wenige, jedoch “schwer zu findende“ Probleme identifizieren konnten und die Streuung der Ergebnisse der Evaluation untereinander, zeigen die Relevanz der Involvierung mehrerer Experten in die Evaluation.

Nielsen spricht in diesem Zusammenhang von drei bis maximal fünf. Mehr als fünf Experten führen nur noch zu einer geringen Zunahme der identifizierten Probleme und demnach zu einer deutlichen Verschlechterung des Kosten-Nutzen-Verhältnisses der Evaluation, siehe Bild 4.14.

Ein einzelner Evaluator erkennt im Schnitt lediglich 35% der Usability-Probleme eines vorhandenen Systems. Die Erhöhung der Anzahl der Evaluatoren steigert auch die relative Anzahl an gefundenen Usability-Problemen, beim Einsatz von fünf Evaluatoren werden bereits 75% der Usability-Probleme identifiziert.

Da die Kurve mit weiter steigender Anzahl der Evaluatoren immer degressiver ansteigt wird das “Kosten-Nutzen“ (Zahl an Evaluatoren/identifizierte Usability Probleme) Ver-

hältnis immer schlechter. Nielsens Empfehlung zur geeigneten Anzahl an Evaluatoren liegt demnach im Bereich von fünf Stück, da diese Anzahl an Evaluatoren gewährleistet, dass ausreichend relevante Probleme in einem Interface identifiziert werden sollten und ein Redesign im Rahmen eines Iterationszyklus sinnvoll ist.

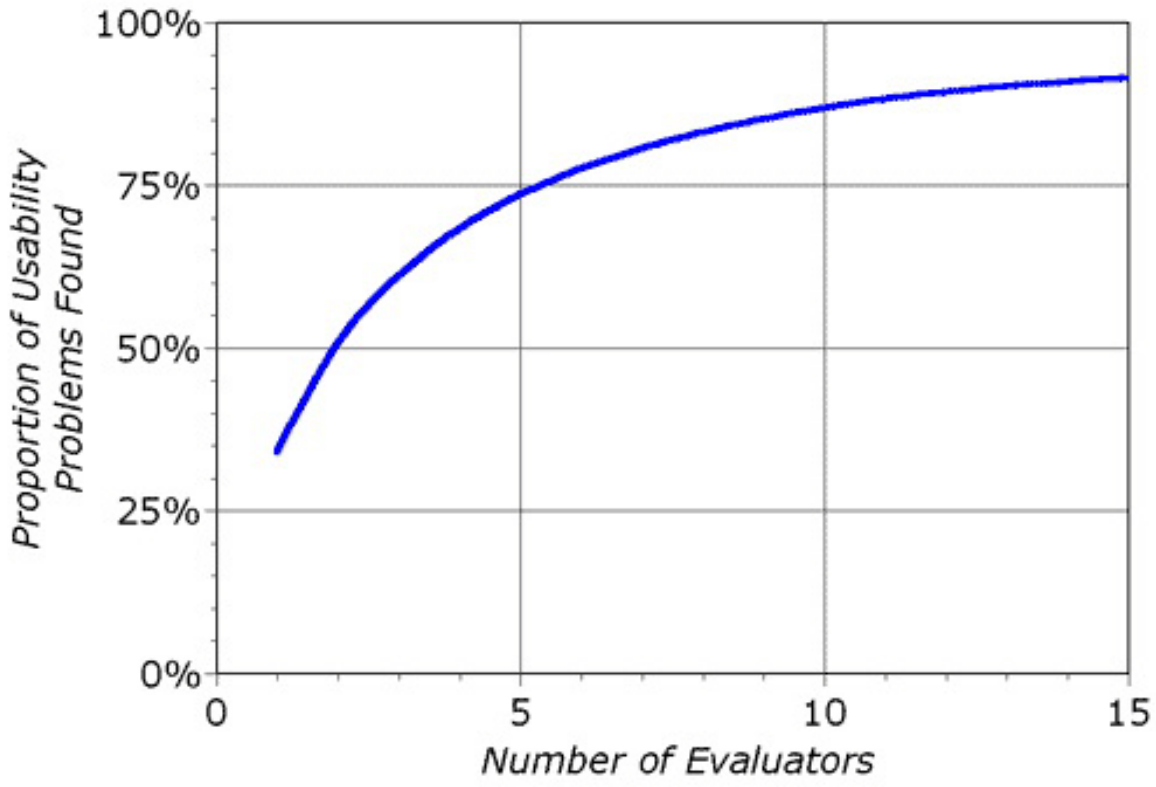

Bild 4.14: Problemfindungskurve Heuristische Evaluation /64/

4.3.4 Abwandlungen und Modifikationen

Es existieren verschiedene Modifikationen der Heuristischen Evaluation, die sich hauptsächlich in der Formulierung und Anzahl der Heuristiken sowie der Partizipation von echten Nutzern unterscheiden..

4.3.4.1 Partizipatorische Heuristische Evaluation:

Die Partizipatorische Heuristische Evaluation ist eine Modifikation der klassischen Heuristischen Evaluation mit einen prozessorientieren Ansatz durch Hinzufügen und Umformulieren neuer und bestehender Heuristiken und Integration von Nutzern als Experten. Tabelle 3 zeigt hierzu als Beispiel die verwendeten Heuristiken der Evaluation einer Spritzenpumpe aus der Medizintechnik:

Nr.	Heuristik	Erläuterung
1.	Steuerbarkeit	Möglichkeit des Anwenders zu jeder Zeit aktiv den Zustand des Systems zu beeinflussen.
2.	Konsistenz	Einheitliche Bedienführung, gleichförmige Gestaltung von Handlungsabläufen(Bediensequenzen).
3.	Erkennen von Bedienfehlern	Fähigkeit des Systems Bedienfehler (z.B. Fehleingaben) zu erkennen und dem Anwender angemessen zurückzumelden.
4.	Prozessunterstützung	Anpassung der erforderlichen Bedienschritte an den zu unterstützenden Arbeitsprozess.
5.	Geräterückmeldungen	Einfache, direkte und verständliche Meldungen und Aufforderungen.
6.	Selbsterklärungsfähigkeit	Intuitive Gerätebedienung, Fähigkeit des Systems den Anwender durch einzelne Bedienschritte zu leiten.

Tabelle 4.3: Beispiel für Heuristiken einer Spritzenpumpe /69/

Sarodnick /35/ schlägt aufgrund der Veralterung der von Nielsen verfassten Heuristiken und einer Integration moderner Ansätze zur Usability-Evaluation von Produkten folgende 12 alternative, generelle Heuristiken vor:

1. **Aufgabenangemessenheit**

Alle benötigten Funktionen für anstehende Aufgaben im System müssen vorhanden sein und hinreichend so gestaltet sein, dass sie den Nutzer unterstützen und ihn bei Routineaufgaben entlasten.

2. **Prozessangemessenheit**

Das System sollte für die Erfüllung realer Arbeitsaufgaben in typischen Einsatzfeldern optimiert sein, einen Bezug zum übergeordneten, realen Prozessziel haben und auf Qualifikationen und Erfahrungen realer Nutzer abgestimmt sein.

3. **Selbstbeschreibungsfähigkeit**

Einheitliche und unmittelbare Anzeige des Systemstatus. Der Benutzer sollte die Detaillierung der Information über den Systemstatus selbst bestimmen.

4. **Steuerbarkeit**

Beinhaltet die Kontrolle des Nutzers über den Dialog, sowie die Möglichkeit, verschiedene Eingabehilfen zu nutzen oder das System ohne Datenverlust zu beenden.

5. **Erwartungskonformität**

Die Informationsdarstellung sollte systemimmanent und mit plattformspezifischen Konzepten konsistent sein. Bei ähnlichen Aufgaben sollten Dialoge vergleichbar und an erwarteter Position dargestellt sein.

6. **Fehlertoleranz**

Fehlermeldungen sollten deutlich sein und Hinweise über Art und Handlungszusammenhang enthalten. Der Nutzer muss über irreversible Handlungen informiert werden.

7. **System- und Datensicherheit**

Das System sollte auch bei fehlerhaften Eingaben des Nutzers und unter hoher Ressourcenbelastung stabil und ohne Datenverluste arbeiten.

8. **Individualisierbarkeit**

Das System sollte sich individuell an die Präferenzen der Nutzer anpassen lassen, solange dies der Effektivität, Effizienz und Zufriedenheit dient und nicht im Widerspruch zu notwendigen, technischen odersicherheitsrelevanten Begrenzungen steht.

9. **Lernförderlichkeit**

Lernstrategien wie “Learning by Doing“ sollten durch schrittweise Anleitungen oder Navigationshilfen unterstützt werden.

10. **Wahrnehmungssteuerung**

Das Layout sollte minimalistisch gehalten werden. Gruppierungen, Farbgestaltung und sinnvolle Informationsreduktion etc. sollten so verwendet werden, dass die Aufmerksamkeit des Nutzers zu relevantem Information gelenkt wird.

11. **Joy of use**

Arbeitsabläufe und graphische Gestaltung des Systems sollten bei notwendiger Konsistenz Monotonie vermeiden und zeitgemäß wirken. Metaphern sollten adäquat und auf den Nutzungskontext abgestimmt verwendet werden.

12. **Interkulturelle Aspekte**

Das System sollte auf einen definierten Nutzerkreis und dessen funktionale, organisatorische und nationale Kultur abgestimmt sein.

4.3.4.2 Kooperative Heuristische Evaluation

Die Kooperative Heuristische Evaluation ist eine weitere Möglichkeit der Modifikation der Heuristischen Evaluation. Der wesentliche Unterschied zum klassischen Verfahren der Heuristischen Evaluation ist die Einbindung von echten Nutzern in den Evaluierungsschritt und die anschließende Kooperation von Experten und Nutzer während der Evaluation.

4.4 Cognitive Walkthrough

Der Cognitive Walkthrough (im folgenden: CWT) ist eine aufgaben-orientierte Usability-Inspektionsmethode, mit dessen Hilfe Experten die Gebrauchstauglichkeit eines technischen Systems oder Interfaces über den Aspekt der schnellen, explorativen Erlernbarkeit durch einen späteren Nutzer, anhand von festgelegten Handlungssequenzen, typischen Arbeitsaufgaben und der Beantwortung von vier Leitfragen, bewerten können.

Die Experten versetzten sich dabei in die Lage der Denkprozesse der späteren Nutzer und versuchen Differenzen zwischen den Annahmen der Entwickler und Designer und dem Wissen der späteren Nutzer aufzudecken. Der CWT analysiert (vorgegebene) korrekte Handlungsabläufe und fragt dabei, ob diesen Abläufen tatsächlich von den Benutzern gefolgt wird /17/. Mit Hilfe des CWT werden folgende drei Problembereiche erfasst:

1. Punkte, an denen die Konzepte der Nutzer und Entwickler über die Aufgaben nicht übereinstimmen.
2. Ungünstige Benennungen der Bedienelemente.
3. Inadäquates Feedback.

Der CWT versucht nicht, die Usability direkt über eine Evaluierung des technischen Systems oder Interfaces zu messen, sondern vielmehr die mentalen Prozesse des späteren Nutzers bei der Bedienung des Systems nachzuvollziehen. Mithilfe des CWT können “kognitive Barrieren“ beim Lern- und Problemlöseprozess des Nutzers identifiziert werden und daraus Schlussfolgerungen über einen erfolgreichen Interaktionsprozess des Nutzers abgeleitet werden. In diesem Sinne fokussiert sich der CWT auf den Gestaltungsgrundsatz der Lernförderlichkeit aus der DIN 9241-110.

Über die Festlegung einer idealen Handlungsabfolge zur Bewältigung eines Arbeitsschritts und die anschließende Analyse, dem “kognitivem Durchgang“ des selbigen, anhand der vier Leitfragen

- Werden die Nutzer versuchen, den gewünschten Effekt zu erzielen?
- Werden die Nutzer erkennen, dass die gewünschte Handlung ausgeführt werden kann?
- Werden die Nutzer erkennen, dass die korrekte Handlung ausgeführt werden kann?
- Werden die Nutzer den Fortschritt erkennen, wenn Sie die korrekte Handlung ausgeführt haben?

werden Aussagen über Erfolg oder Misserfolg der Handlung gewonnen. Ziel des CWT ist die Ermittlung der leichten Erlernbarkeit der Bedienung eines Interfaces eines interaktiven Systems.

4.4.1 Theorie

Grundlage des CWT ist die Kognitionsforschung und die CE+-Theorie /29/ des explorativen Lernens, deren Modell aus folgenden Komponenten besteht:

- Problemlösekomponente (problem-solving-component)
- Lernkomponente (learning component)
- Ausführungskomponente (execution component)

Der allgemeine Problemlöseprozess eines Nutzers beim explorativen Erkunden und Erlernen eines Interfaces und das Verhalten, dass entweder zum Erfolg oder Misserfolg führen kann unterliegt folgenden Annahmen /31/:

- The user sets a goal to be accomplished with the system (for example, "check spelling of this document").
- The user searches the interface for currently available actions (menu items, buttons, command-line inputs, etc.).
- The user selects the action that seems likely to make progress toward the goal.
- The user performs the selected action and evaluates the system's feedback for evidence that progress is being made toward the current goal.

Sarodnick /36/ formuliert diese Annahmen wie folgt:

- Nutzer gehen von einer Grobbeschreibung der Aufgabe aus, die sie mit dem System bearbeiten wollen.
- Sie ergründen eine Schnittstelle und wählen dann Handlungen, die direkt oder über Zwischenschritte zur Lösung führen werden.
- Sie beobachten die Systemreaktionen, um herauszufinden, ob die ausgeführten Aktionen den gewünschten Effekt haben.
- Sie legen fest, welches der nächste Handlungsschritt ist.

Aufbauend auf diesem Modell wurde in der ACM 1990 die erste Version des CWT von Lewis et al. Veröffentlicht, siehe Bild 4.15. In einer Studie /17/ wurde festgestellt dass 50% der Usability-Probleme, die durch eine umfangreiche, empirische Vergleichsevaluation festgestellt wurden, mit dem CWT gefunden wurden, mit minimalen Investitionskosten in Zeit, Geld und Ressourcen bei einer Bearbeitungsdauer von einer Stunde pro Einzeltätigkeit und Interface. Es wurde jedoch auch festgestellt, dass die Evaluationsergebnisse von dem Theoriewissen der Evaluatoren beeinflusst wurden und die erfolgreiche Durchführung der Methode nur durch Aneignung und Training von Grundlagenwissen und der CE+-Theorie möglich sei.

Die zweite Version des CWT /28/ war wesentlich detaillierter beschrieben in Bezug auf das CE+-Modell, wie auch den CWT selbst und sollte auch Anwendern ohne kognitionswissenschaftlichen Hintergrund die korrekte Durchführung eines CWT ermöglichen,

siehe Bild 4.16. Die Fragen der Analysephase wurden von 19 auf 17 Stück reduziert, und die Analysephase in drei statt in neun Teile zusammengefasst.

In der aktuellen Version des CWT sind jedoch keine kognitionswissenschaftlichen Grundkenntnisse mehr nötig. Die Methode kann auch von Laien ohne detaillierte Vorkenntnisse verstanden und angewandt werden. Das grundlegende Vorziehen des explorativen, intuitiven Lernens bei der Bedienung eines Interfaces durch den Benutzer, statt formales Training, wie beispielsweise durch Lesen einer Bedienungsanleitung oder Benutzung einer Hilfe, ist der eigentliche Ansatz der Methodik des CWT. Benötigt der Benutzer zur Bedienung des Interfaces Wissen, dass er nicht besitzt oder verlangt das System dem Benutzer Wissen ab, das er nicht hat, so kann die Usability nicht besonders hoch sein.

CE+ Design Walkthrough **Evaluator** _____ **Date** ____

Interface _____ **Task** _____ **Step #** ___

Actions/choices should be ranked according to what percentage of potential users are expected to have problems: 0 = none; 1 = some; 2 = more than half ; 3 = most.

1. Description of user's immediate *goal:*
2. (First/next) atomic *action* user should take:
   - 2a. Obvious that action is *available* ? Why/Why not?
   - 2b. Obvious that action is *appropriate to goal* ? Why/Why not?
3. How will user *access description* of action?
   - 3a. Problem accessing? Why/Why not?
4. How will user *associate* description with action?
   - 4a. Problem associating? Why/why not?
5. All other available actions *less appropriate* ? For each, why/why not?
6. How will user *execute* the action?
   - 6a. Problems? Why/why not?
7. If *timeouts,* time for user to decide before timeout? Why/why not?
8. Execute the action. Describe *system response:*
   - 8a. Obvious *progress* has been made toward goal? Why/why not?
   - 8b. User can access needed *info. in system response?* Why/why not?
9. Describe appropriate *modified goal,* if any:
   - 9a. Obvious that *goal should change?* Why/why not?
   - 9b. If task *completed,* is it obvious? Why/why not?

Figure 1. A cognitive walkthrough evaluation form for a single action

Bild 4.15: Erste CWT Anleitung für einen einzelnen Arbeitsschritt /65/

Task:_____ **Action#**_____

1. Goal structure for this step

1.1 **Correct goals.** What are the appropriate goals for this point in the interaction?

1.2 **Mismatch with likely goals.** What percentage of users will not have these goals, based on the analysis at the end of the previous step?

2. Choosing and executing the action. Correct action at this step:_____

2.1 **Availability.** Is it obvious that the correct action is a possible choice here?

2.2 **Label.** What label or description is associated with the correct action?

2.3 **Link of label to action.** If there is a label or description associated with the correct action, is it obvious, and is it clearly linked with this action?

2.4 **Link of label to goal.** If there is a label or description associated with the correct action, is it obviously connected with one of the current goals for this step?

2.5 **No label.** If there is no label associated with the correct action, how will users relate this action to a current goal?

2.6 **Wrong choices.** Are there other actions that might seem appropriate to some current goal?

2.7 **Time-out.** If there is a time-out in the interface at this step does it allow time for the user to select the appropriate action?

2.8 **Hard to do.** Is there anything physically tricky about executing the action?

3. Modification of goal structure. If the correct action has been taken, what is the system response?

3.1 **Quit or backup.** Will users see that they have made progress towards some current goal?

3.2 **Accomplish goals.** List all current goals that have been accomplished. Is it obvious that they have been accomplished?

3.3 **Incomplete goals that look accomplished.** Are there any current goals that have not been accomplished, but might appear to have been based on the system response?

3.4 **"And-then" structures.** Is there any "and-then" structure, and does one of its subgoals appear to be complete?

3.5 **New goals in response to prompts.** Does the system response contain a prompt or cue that suggests any new goal or goals?

3.6 **Other new goals.** Are there any new goals that users will form given their current goals, the state of the interface, and their background knowledge?

Bild 4.16: CWT, zweite Version, Analysephase /28/

4.4.2 Detaillierter Ablauf des Verfahrens

Der CWT besteht aus einer Vorbereitungsphase und einer Analysephase.

Vorbereitungsphase

In der Vorbereitungsphase werden durch die Evaluatoren die Grundlagen und der Input des CWT festgelegt. Dies geschieht über die Definition und Dokumentation folgender vier Bereiche /32/:

- Annahmen über die Nutzer/Zielgruppe
- Festlegen der zu analysierenden Aufgaben
- Festlegen der idealen Handlungssequenz zur erfolgreichen Bewältigung einer Aufgabe
- Definition der Schnittstellen/Beschreibung des Interfaces

Um in der späteren Analysephase Differenzen zwischen dem Vorwissen der Nutzer und den erforderlichen Kenntnissen zur Bedienung des Systems erkennen zu können, müssen zuerst die Charakteristika der Nutzer und deren Vorwissen ermittelt werden. Um ein

prototypisches, möglichst repräsentatives Nutzerprofil generieren zu können, werden zuerst allgemeine Daten über die Eigenschaften der zukünftigen Nutzer des technischen Systems definiert, beispielsweise Alter, Geschlecht und Vorkenntnisse im Umgang mit dem Produkt.

Das Festlegen der zu analysierenden Aufgaben ist ein kritischer Punkt in der Vorbereitungsphase. Aufgrund der Komplexität der meisten technischen Systeme können in der Praxis nicht alle Aufgaben berücksichtigt werden. Daher werden durch die Evaluatoren spezifische Einzelaufgaben ausgewählt, die möglichst bedeutsam und realistisch für den alltäglichen Arbeitsprozess des späteren Nutzers sind. Eine Ableitung der Aufgaben aus einem Pool von möglichen Anwendungs- und Aufgabenszenarien des technischen System kann die Auswahl erleichtern, es sollten in jedem Fall häufig vorkommende, kritisch- realistische und repräsentative Aufgaben gewählt werden.

Anschließend wird eine ideale, korrekte Bediensequenz zu jeder Beispielaufgabe definiert, beschrieben und dokumentiert. Hierbei kann die Miteinbeziehung von Entwicklern und Designern nützlich sein. Sollten mehrere Lösungsalternativen zur Verfügung stehen, wird entweder die gebräuchlichste oder die kritischste Aktionsabfolge gewählt.

In dem letzten Schritt der Vorbereitungsphase wird detailliert beschrieben, was der Nutzer bei dem jeweiligen Handlungsabschnitt zu sehen bekommt, beispielsweise die Anzahl und Position der Auswahlbuttons, deren Farbe und Form usw. Ist die Schnittstelle noch nicht als Prototyp verfügbar, kann dies schriftlich erfolgen, ansonsten bieten sich erläuterte Screenshots, Fotos oder Mock-Up-Screens an.

Analysephase

In der Analysephase wird der eigentliche CWT durch die Experten gemeinsam durchgeführt, d.h. die Experten versetzten sich in die hypothetische Nutzer und gehen chronologisch die Einzelaufgaben, Schritt für Schritt, einer, ggf. gesplitteten Gesamtaufgabe, durch. Anhand der eingeführten vier Leitfragen, wird jeder Schritt innerhalb einer Handlungssequenz untersucht und eine Erfolgs- oder Misserfolgsstories abgeleitet. Wird nur eine dieser Fragen mit “Nein“ beantwortet folgt daraus eine “Failure-Story“, das bedeutet der Nutzer kann die Aktion nicht erfolgreich durchführen. Nur die Beantwortung aller vier Leitfragen mit “Ja“ führt demnach zu einer “Success-Story“.

Im Folgenden werden die vier Leitfragen des CWT zur Bewertung einer erfolgreichen Nutzerinteraktion vorgestellt und erläutert:

1. **Werden die Nutzer versuchen, den gewünschten Effekt zu erzielen?**

Ob der Benutzer überhaupt die Assoziation zwischen gewünschten Ziel und der korrekten Aktion, die zur Erreichung des selbigen vonnöten ist, herstellen kann, wird durch

diese Frage beantwortet. Nutzer müssen sich ihres Ziels bewusst sein und erkennen können, welche Handlungen zur Erfüllung notwendig sind. Dies ist der Fall wenn:

- sie bereits Erfahrung mit dem System haben,
- das System sie zur korrekten Handlung auffordert ,
- diese Handlungen bereits Teil ihrer ursprünglichen Arbeit sind.

2. **Werden die Nutzer erkennen, dass die korrekte Handlung ausgeführt werden kann?**

Wenn der Nutzer nun weiß, dass er eine bestimmte Handlung ausführen muss, um einen Effekt oder ein Ziel zu erreichen, so muss ihm das System auch aufzeigen, dass er diesen Schritt auch ausführen kann. Hier geht es zum Beispiel um eine visuelle Identifikation eines Auswahlbuttons oder eines Menüs, zum benötigten Zeitpunkt der Handlungssequenz.

3. **Werden die Nutzer erkennen, dass die korrekte Handlung zum gewünschten Effekt führen wird?**

Der Benutzer kennt nun die korrekte Handlung, die zum gewünschten Effekt führt, ist jedoch nicht in der Lage die Aktion mit dem gewünschten Effekt in Verbindung zu bringen. Analog zu dem vorherigen Beispiel bedeutet dies, dass der Nutzer den benötigten Auswahlbutton oder das Menü sieht, sich aber nicht sicher ist, ob die Auswahl dieses einen Elements zum gewünschten Effekt führt. Nutzer können die Verbindung zwischen Handlung und gewünschten Effekt herstellen, wenn:

- sie bereits Erfahrung mit dem System haben,
- das System auf die Verbindung hinweist,
- alle anderen Handlungen weniger Erfolg versprechend erscheinen.

4. **Werden die Nutzer den Fortschritt erkennen, wenn sie die korrekte Handlung ausgeführt haben?**

Hier ist die Frage, ob der Benutzer, nachdem er die korrekte Aktion getätigt hat, angemessenes Feedback vom System erhält und ihm der Erfolg seiner Handlung mitgeteilt wird. Ein Beispiel hierfür wäre eine Sendebestätigung eines Emailprogramms beim Verschicken einer email. Nutzer erkennen den Fortschritt, wenn:

- sie bereits Erfahrung mit dem System haben,
- eine Systemreaktion stattfindet und der Nutzer diese in Verbindung mit seiner ausgeführten Handlung bringen kann.

Diese Handlungsabfolge ist in jedem Fall einzuhalten. Eventuelle “Failure-Stories“ bei der Beantwortung einer Frage, werden für die nachfolgenden Teilschritte als erfolgreich

angenommen. Bild 4.17 zeigt exemplarisch einen Arbeitsbogen für einen CWT einer Website.

Arbeitsbogen für einen Cognitive Walkthrough durch eine Website

URL: **Titel:**

Beispielaufgabe: [z.B.: Simulation xy finden und durchspielen]

Was fällt auf? Beschreiben Sie, was Sie auf der Eingangsseite spontan sehen:

Idealer Lösungsweg für den Nutzungsfall (Use Case)

1. Schritt: [tragen Sie hier den korrekten Schritt ein, z.B. „Das „pull down Menü" finden, draufklicken, die Eintragungen verstehen, auf ‚Kopieren' klicken"]

Success-Story

☐ Der Anwender wird versuchen, die korrekte Aktion auszulösen.
☐ Der Anwender kann erkennen, dass die gewünschte Aktion zur Verfügung steht.
☐ Der Anwender kann den Zusammenhang zwischen der korrekten Aktion und der gewünschten Auswirkung erkennen.
☐ Der Anwender sieht einen Erfolg seiner Handlung.

Failure-Story

☐ Der Anwender wird nicht versuchen, die korrekte Aktion auszulösen.
☐ Der Anwender kann nicht erkennen, dass die gewünschte Aktion zur Verfügung steht.
☐ Der Anwender kann den Zusammenhang zwischen der korrekten Aktion und der gewünschten Auswirkung nicht erkennen.
☐ Der Anwender sieht keinen Erfolg seiner Handlung.

Was fällt auf? Beschreiben Sie, was Sie auf der neuen Seite nach dem 1. Schritt sehen:

Bild 4.17: Beispielbogen eines CWT für eine Website /67/

4.4.3 Dokumentation des CWT

Die Evaluationsbögen sollten als Kopfinformationen folgendes beinhalten:

- Eine Beschreibung des Interfaceprototyps.
- Eine Beschreibung der Arbeitsaufgabe des Nutzers innerhalb dieses Systems.
- Eine Liste der Handlungen, die der Nutzer für eine erfolgreiche Bewältigung seiner Arbeitsaufgabe tätigen muss.
- Eine Grobbeschreibung der Nutzer und deren Hintergrundwissen und Erfahrung mit dem System.

Für jede Einzelhandlung innerhalb der Sequenz wird einen separater Evaluationsbogen benutzt, der die Beantwortung der vier Leitfragen beinhaltet. Eventuelle Misserfolgsstories werden auf einen weiteren Berichtsbogen dokumentiert.

4.4.4 Lösungsfindungsprozess/Revision des Interfaces

Die “Failure-Stories“ stehen in direktem Zusammenhang mit den vier Leitfragen, daher können sie auch zur Lösungsfindung und Revision des Interfaces benutzt werden:

1. **Der Benutzer versucht nicht, den richtigen Effekt zu erzielen**

   Mögliche Lösungsvorschläge:

   - Automatisierung der Handlung durch das System oder Kombination mit anderer Handlung.
   - Hinweis geben, welches die nächste notwendige Aktion ist.
   - Durch Umgestaltung der Aufgabe klareres Herausstellen der benötigten Aktion.

2. **Der Nutzer erkennt nicht, dass die korrekte Handlung ausgeführt werden kann**

   Möglicher Lösungsvorschlag:

   - Verwendung offensichtlicherer Bedienelemente als vorher.

3. **Der Nutzer erkennt nicht, dass die korrekte Handlung zum gewünschten Effekt führen wird.**

   Möglicher Lösungsvorschlag:

   - Verbesserung der Anordnung und Gestaltung der Bedienelemente in Bezug auf die Benutzer und deren Aufgabenverständnis. Beschriftung anhand realistischer Aufgabenelemente ausrichten.

4. **Der Nutzer erkennt nicht den Fortschritt, der durch die korrekte Handlung entsteht.**

Möglicher Lösungsvorschlag:

- Optimierung des Feedbacks des Systems.

4.4.5 Streamlined Cognitive Walkthrough

Die aktuellste Version des Cognitive Walkthrough, der Streamlined CWT /40/, ist eine reduzierte Version des in Kapitel 4.4.2 beschriebenen klassischen Verfahrens. Erfahrungen in der Praxis mit der klassischen Methode haben gezeigt, dass der CWT von Entwicklern oft als zu langsam, schwerfällig und redundant empfunden wurde.

Die wesentliche Neuerung ist die Reduktion der ersten drei Fragen des CWT auf eine, so dass insgesamt nur noch zwei Fragen innerhalb eines Aktionsschrittes beantwortet werden müssen:

1. **Wird der Benutzer wissen, was er an dieser Stelle tun muss?**

2. **Sollte der Benutzer das Richtige getan haben, wird er wissen, dass er das Richtige getan hat, und dass er einen Fortschritt in Richtung seines Ziels gemacht hat?**

Abgesehen von dieser Änderung entspricht die Vorgehensweise dem beschriebenen CWT.

4.5 KLM-GOMS

Unter dem GOMS-Modell versteht man ein analytisches Verfahren, dessen Ziel die Modellierung zielgerichteten, menschlichen Handelns ist und zur Vorhersage von Ausführungszeiten für bestimmte Handlungen in Bildschirmsystemen benutzt werden kann. Das Akronym GOMS /47/ steht für **G**oals, **O**perators, **M**ethods, **A**ction Sequences.

Der Grundgedanke des GOMS-Modells ist die Annahme, dass Benutzer in einem interaktiven System ihre globalen Ziele und Aufgaben in voneinander unabhängige Teilziele und Aufgaben unterteilen, um diese zu erreichen und zu lösen.

Als theoretisches Konzept zur Modellierung menschlichen Handelns innerhalb der Mensch-Computer-Interaktion gibt das Modell Zeiten vor, die für bestimmte Handlungen und Kognitionen zur Aufgabenbearbeitung benötigt werden. Über eine anschließende Kopplung und Summation der theoretisch benötigten Zeit der Einzelschritte, lässt sich eine Gesamtzeit zur Aufgabenbearbeitung errechnen, was wiederum zur vergleichenden Bewertung der Effizienz von Interfacegestaltungsentwürfen verwendet werden kann.

Stehen zum Beispiel verschiedene Konzepte eines Interfacedesigns mit gleicher Funktionalität zur Verfügung, kann mit dem KLM-GOMS-Modell innerhalb jedes Gestaltungsentwurfs, eine Zeit berechnet werden, die für die Bearbeitung einer definierten Aufgabe theoretisch benötigt wird. Anschließend lassen sich über einen Vergleich der theoretisch berechneten Arbeitszeiten Aussagen über die Effizienz der verschiedenen Interfacelayouts ableiten.

4.5.1 Die vier Komponenten des GOMS-Modells

Die angenommene Ausgangssituation innerhalb des GOMS-Modells geht von einem Benutzer aus, der ein globales Gesamtziel hat, dass mit dem gegebenen interaktiven System lösbar ist. Das Bildschirmsystem stellt dem Benutzer bestimmte Operatoren zur Aufgabenbewältigung zur Verfügung. Der Benutzer wendet Sequenzen von notwendigen Operatoren zur Aufgabenbewältigung an und selektiert diese selbstständig hinsichtlich ihrer Eignung zur Aufgabenbewältigung.

Ziele

Die Ziele sind hierarchisch gegliedert, d.h. es existiert nur ein globales Gesamtziel, z.B. “Email verschicken“ oder “Dokument drucken“. Zur Zielerreichung teilt der Benutzer das globale Gesamtziel in Teilziele auf, zum Beispiel Email Programm öffnen, Email schreiben und Email versenden. Der Benutzer arbeitet die Ziele und Teilziele sequentiell ab, eine parallele Bearbeitung der Teilziele wird ausgeschlossen.

Operatoren

Die Operatoren sind die elementaren Handlungen des Benutzers zur Zielerreichung. Unter den Operatoren versteht man kognitive Handlungen, wie mentale Vorbereitung aber auch motorische Bewegungen, wie Maustaste klicken und Wahrnehmung des Systemzustandes. Der gesamte Handlungsablauf des Benutzers zur Zielerreichung wird innerhalb des GOMS-Modells auf die Operatoren reduziert, sein Verhalten also modelliert und anschließend formalisiert.

Operatoren benötigen Eingaben durch den Benutzer und haben eine entsprechende Wirkung, erzeugen entsprechende Ausgaben und verbrauchen Zeit. Innerhalb des Keystroke-Level-Modells, das zur Berechnung der notwendigen Bearbeitungszeit über die Gesamtzahl von Tastendrücken und Mausklicks verwendet wird, existieren folgende Operatoren:

Operator	K	T(n)	P	B	BB	H	M	W(t)
Zeit [sec]	0,28	n*K	1,1	0,1	0,2	0.4	1,2	n.b.

Tabelle 4.4: Operatoren des KLM-GOMS Modells

K: Keystroke

Die Zeit zum Drücken einer Taste auf dem Keyboard. Der Erfahrungslevel des Benutzers bei der Arbeit mit der Tastatur wird durch folgende Abstufung berücksichtigt:

Erfahrungslevel des Nutzers	**Zeit [sec] pro Anschlag**
Experte	0,12
Überdurchschnittliche Erfahrung mit der Tastatur	0,20
Durchschnittsnutzer	0,28
Keine Erfahrung mit einer Tastatur	1,2

Tabelle 4.5: Zeit pro Anschlag im KLM-GOMS-Modell

T(n): Sequenz

Die Zeit einer Sequenz von n Tastenanschlägen. Der Operator dient lediglich der Zusammenfassung von einer Serie von Zeichenketten die mit der Tastatur eingeben werden.

P: Pointing

Die Zeit zum Positionieren einer Maus auf einem Punkt des Bildschirms.

B: Drücken oder Loslassen der Maustaste

Die Zeit die benötigt wird um einen Mausklick durchzuführen oder die Maustaste loszulassen.

BB: Klicken der Maustaste

Drücken und sofortiges Loslassen der Maustaste, also ein einfacher Klick.

H: Homing

Die Zeit, die benötigt wird, um mit der Hand von dem Keyboard zur Maus zu wechseln und umgekehrt.

M: Mentally preparing

Die Zeit zur mentalen Vorbereitung des nächsten Arbeitsschrittes.

W(t):

Die Zeit, die der Benutzer nach einer Eingabe auf eine Systemreaktion warten muss.

In der Anwendung wird die Handlungsabfolge des Benutzers in einem Computersystem durch die Operatoren formal beschrieben. Für die Verwendung des Mentaloperators, der die Nachdenkzeit des Nutzers für die nächstmögliche Handlung symbolisiert, existieren definierte Regeln /47/.

Methoden

Die Methoden sind Sequenzen von Operatoren zur Erreichung eines Teilziels oder des Gesamtziel.

Auswahlregeln

Die Auswahlregeln sind Wenn-Dann Beziehungen, die entscheiden, welches Handlungsschemata bei verschiedenen Möglichkeiten zur Zielerreichung verwendet wird.

4.5.2 Einsatzbereich des KLM-GOMS Modells

Das KLM-GOMS Modell eignet sich nur für einen sehr beschränkten Einsatzbereich im Vergleich zu den anderen vorgestellten Methoden zur Usability-Evaluation. Mit den

theoretisch berechneten Ausführungszeiten lassen sich Gestaltungsalternativen für Computersoftware hinsichtlich ihrer Effizienz vergleichen.

Stehen zum Beispiel in einem Softwareprogramm drei Befehle und Aktionen zum Löschen einer Datei zur Verfügung, können diese Alternativen zeitlich verglichen werden.

Im vorgestellten KLM-GOMS Modell, das als sehr einfaches und schnelles Verfahren zur Bewertung von Gestaltungsalternativen in Bildschirmsystemen benutzt werden kann, werden ausschließlich die Operatoren zur Formalisierung der Zeitspanne, die eine Nutzerinteraktion voraussichtlich dauert, benutzt. Die Auswahlregeln werden in dieser einfachsten Variante des GOMS-Modells nicht berücksichtigt.

5 Bewertung

Die hier vorgestellten Verfahren zur Evaluation stellen eine Auswahl der am häufigsten in der Literatur vorgefundenen und in der Praxis angewandten Methoden zur Usability-Evaluation dar.

Der direkte Vergleich der Evaluationsmethoden ist aufgrund ihrer völlig heterogenen Ansätze in der Erhebung sehr schwer. Die Methoden unterscheiden sich vor allem in ihrer Fragestellung und demnach auch ihrer Zielsetzung, der Art der Datenerhebung, dem letztlich erhobenen Datentyp und der Beteiligung und Nichtbeteiligung von Nutzern und/oder Experten.

Im Folgenden werden Bewertungskriterien vorgestellt, anhand derer, eine erste, grobe Einschätzung der Ausprägung verschiedener Aspekte der vorgestellten Methoden zur Usability-Evaluation möglich ist. Die Kriterien, sowie deren Bewertung, wurden dem für diese Studienarbeit verwendeten Grundlagenbuch “Methoden der Usability-Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung“ von Florian Sarodnick /37/ entnommen.

Durch die Einführung zusätzlicher Kriterien wie dem Standardisierungsgrad, der Bandbreite der erhobenen Daten und der Anwendbarkeit im Entwicklungsprozess, soll eine noch differenziertere Betrachtung der vorgestellten Methoden ermöglicht werden.

Die Ausprägung der einzelnen Kriterien wird durch eine Punkteskala von ● bis ●●● bewertet. Nicht bewertbare Kriterien, wie beispielsweise die Produktivität der Fragebogenverfahren werden entsprechend mit einem “n.b.“ gekennzeichnet.

Vergebene Punktezahl	**Ausprägung des Kriteriums**
●	Gering
●●	Mittel
●●●	Hoch
n.b.	Nicht bewertbar

Tabelle 5.1: Legende zur Bewertung

5.1 Bewertungskriterien

5.1.1 Kriterien mit praktischer Relevanz

Produktivität:

Anzahl der identifizierten Usability-Probleme bei Anwendung der Methode.

materieller Aufwand:

Kosten für Vorbereitung, Durchführung und Auswertung sowie Raumbedarf, die benötigte Hard- und Software und sonstige Auswertungswerkzeuge. Personelle Kosten werden hier ebenfalls berücksichtigt. Dazu zählen, abhängig von der Methode, Benutzer, Experten, Versuchsleiter, Softwaredesigner und Techniker.

zeitlicher Aufwand:

Gesamtzeit für Planung, Vorbereitung, Durchführung und Auswertung der Evaluation.

Notwendige Qualifikation der Evaluatoren:

Ausmaß der benötigten Fachkenntnisse und Usability-Expertise der Evaluatoren, die bei Anwendung der Methode benötigt werden.

Detaillierungsgrad:

Aussagekraft und Konkretisierungsgrad der erhobenen Daten durch die Evaluation. Unter einem hohen Detaillierungsgrad wird zum Beispiel die Erhebung von exakten objektiven Daten wie Ausführungszeiten oder Fehlerraten sowie subjektive Daten aus erster Hand wie echten Benutzermeinungen verstanden. Der Detaillierungsgrad hängt somit eng mit dem jeweiligen Output der Methode zusammen.

Flexibilität der Methode:

Anpassungsfähigkeit und Variationsvielfalt der Methode für verschiedene Fragestellungen. Stark formalisierte Verfahren wie Fragebögen und KLM-GOMS sind in der Anwendung kaum flexibel. Dem hingegen sind Usability-Tests in der Datenerhebung und Fragestellung variabel und können für nahezu alle möglichen Fragestellungen verwendet werden.

Identifikation individueller Nutzungsprobleme:

Ausmaß der Identifikation von individuellen Nutzungsproblemen. Formalisierte Verfahren, die eine generelle, globale Bewertung der Usability anhand von festgelegten, stan-

dardisierten Vorgehensweisen vornehmen, werden im Vergleich zu anpassungsfähigen, Benutzer- und Experteninvolvierenden Verfahren schlechter bewertet.

Anwendbarkeit im Entwicklungsprozess:

Eignung der Methode in den verschiedenen Phasen des Entwicklungs- und Konstruktionsprozesses. Aufgrund der grundlegend verschiedenen Ansätze der Verfahren und ihren spezifischen Fragestellungen und Zielsetzungen zur Bewertung der Usability, sind manche nur für bestimmte Phasen im Entwicklungsprozess geeignet.

Standardisierungsgrad:

Verfahren, denen eine eindeutig definierte Vorgehensweise zugrunde liegt, besitzen einen hohen Standardisierungsgrad. Ein hoher Standardisierungsgrad hat entsprechend eine niedrige Flexibilität zur Folge.

Bandbreite der erhobenen Daten

Die Bandbreite der erhobenen Daten wird durch die Methodenauswahl und das Erhebungsverfahren determiniert.

5.1.2 Kriterien mit wissenschaftlicher Relevanz

Vorhersagekraft:

Relevanz der identifizierten Probleme der Methode für den tatsächlichen Endanwender. Die Methoden unterscheiden sich weiterhin in der Beteiligung/Nichtbeteiligung von repräsentativen Nutzern und Experten. Methoden mit hoher Vorhersagekraft binden Nutzer in die Evaluation mit ein. Die formal-analytischen Verfahren, die theoretische Probleme ohne empirische Überprüfung feststellen, werden mit einer geringeren Vorhersagekraft bewertet.

Evaluator-Effekt:

Unabhängigkeit der Ergebnisse von dem jeweiligen Evaluator.

Objektivität:

Unabhängigkeit der Ergebnisse vom Versuchsleiter.

Reliabilität:

Ausmaß der Ergebnisübereinstimmung von Evaluationen desselben Systems mit derselben Methode.

5.2 Bewertungstabelle

	Benutzertest[1]	**Fragebogen**	**Heuristische Evaluation**	**Cognitive Walkthrough**	**KLM-GOMS**
Erhobener Datentyp	Subjektiv & Objektiv	Subjektiv	Subjektiv	Subjektiv	Objektiv
Output der Methode	Hinweisliste zu konkreten Problemen von realen Nutzern; objektive Performance-Messung.	Quantitative Auswertung von Benutzermeinungen zu gegebenen Fragestellungen; Hinweise auf Problembereiche.	Gewichtete Hinweisliste und Kategorisierung zu gefundenen Usability-Problemen.	Erfolgs und Misserfolgsstories über die Interaktion des Benutzers, ggf. Ableitung von Lösungsvorschläge.	Absolut-Angaben der Bearbeitungszeit für eine Handlung oder Aktion in einem interaktiven System.
Kriterien mit Praxisrelevanz					
Produktivität	●●●	**n.b.**	●●●	●●	●
Materieller Aufwand	●●●	●	●●	●●	●
Zeitlicher Aufwand	●●●	●	●	●●	●
Qualifikation der Evaluatoren	●●●[2]	●	●●	●●	●
Detaillierungsgrad	●●●[3]	●	●●●[3]	●●[3]	●●●
Flexibilität	●●●	●	●●●	●●●	●
Identifikation individueller Nutzungsprobleme	●●●	●	●●	●●	●
Anwendbarkeit im Entwicklungsprozess	●●	●	●●	●●●	●

Tabelle 5.2: Bewertung der Methoden

Tabelle 5.2: Bewertung der Methoden - Fortsetzung

	Benutzertest[1]	**Fragebogen**	**Heuristische Evaluation**	**Cognitive Walkthrough**	**KLM-GOMS**
Erhobener Datentyp	Subjektiv & Objektiv	Subjektiv	Subjektiv	Subjektiv	Objektiv
Output der Methode	Hinweisliste zu konkreten Problemen von realen Nutzern; objektive Performance-Messung.	Quantitative Auswertung von Benutzermeinungen zu gegebenen Fragestellungen; Hinweise auf Problembereiche.	Gewichtete Hinweisliste und Kategorisierung zu gefundenen Usability-Problemen.	Erfolgs und Misserfolgsstories über die Interaktion des Benutzers, ggf. Ableitung von Lösungsvorschläge.	Absolut-Angaben der Bearbeitungszeit für eine Handlung oder Aktion in einem interaktiven System.
Kriterien mit praktischer Relevanz					
Standardisierungsgrad	●	●●●	●●●	●●	●●●
Bandbreite der erhobenen Daten	●●●	●●	●●	●	●
Kriterien mit wissenschaftlicher Relevanz					
Vorhersagekraft	●●●	●●●	●●	●●	●
Evaluator-Effekt	●●●	●●[4]	●●●[4]	●	●
Objektivität	●●	●●●	●	●	●●●
Reliabilität	●●	●●●	●	●●	●●●

[1] Es wird von mindestens fünf teilnehmenden Nutzern ausgegangen.

[2] Entspricht der Qualifikation des Versuchsleiters.

[3] Da keine formalisierte Fehlerbeschreibung vorliegt, stark abhängig von Evaluator und Erhebungsmaterial.

[4]Bezogen auf die Ergebnisse der einzelnen Evaluatoren – der Output besteht aus einem Gruppenkonsens.

6 Theoretische Beschreibung der Methoden am einem Beispiel

Im folgenden Kapitel werden anhand eines typischen Consumer-Produkts, einer handelsüblichen Digitalkamera mittlerer Preisklasse, Aufgabenszenarien und Zielgruppendefinitionen exemplarisch formuliert und die Durchführung der vorgestellten Methoden theoretisch an diesem Beispiel erläutert.

6.1 Produktdefinition - Zielgruppendefinition

Bevor die Evaluationen durchgeführt werden, muss definiert werden welches Produkt evaluiert wird und in welchen Zustand sich das Evaluationsobjekt befindet.

Betrachtet wird eine Digitalkamera der Firma Canon, Modellbezeichnung Digital Ixus 70 /48/.

Übersicht der Stellteile und Anzeigen:

Bild 6.1: Stellteile und Anzeigen – Kameraoberseite

Bild 6.2: Stellteile und Anzeigen – Kamerarückseite

6.1.1 Definition der Ausgangssituation vor der Evaluation

Die Kamera ist voll aufgeladen, mit einer handelsüblichen zwei Gigabyte SD-Speicherkarte bestückt und befindet sich im ausgeschalteten Zustand.

Das Gerät ist nach dem Einschalten einsatzbereit und die Dokumentation der Kamera ist verfügbar und liegt bereit.

Es existiert durch Herstellerseite bereits eine Zielgruppendefinition, die Merkmale und Kennzeichen wie Alter, Geschlecht, Vorkenntnisse der Benutzer in Bezug auf Kamera und Digitalfotografie der zu erwartenden Nutzergruppe enthält. Diese Zielgruppendefinition liegt für die verschiedenen Evaluationsverfahren vor und wird auch zur Nutzungskontextdefinition benutzt.

6.1.2 Definition des Evaluationsziels

Im nächsten Schritt werden die Ziele der Evaluation formuliert (siehe Kapitel 4.2.1). Dem Auftraggeber sind folgende Fragestellungen wichtig:

- Ist die Kamera einfach und schnell zu bedienen?
- Gibt es Schwachstellen bei der Bedienung?
- Wo gibt es Verbesserungspotential?
- Wie kommt das Produkt bei den Kunden an?

Da es sich bei dem untersuchten Produkt um kein Computersystem handelt und zudem kein Vergleich von Ausführungszeiten für Gestaltungsalternativen erhoben werden soll, eignet sich das vorgestellte KLM-GOMS Verfahren nicht zur Evaluierung in diesem Szenario. Die Heuristische Evaluation, der Benutzertest in Verbindung mit einem Fragebogenverfahren sowie der CWT können für diese Zielsetzung angewendet werden, wobei die Messung der Akzeptanz und Zufriedenheit beim Umgang mit der Digitalkamera nur im Benutzertest nachgewiesen werden kann.

6.1.3 Ableitung von Arbeitsaufgaben aus den Kernfunktionen

Das Handbuch der Digitalkamera /49/ beschreibt auf 148 Seiten die Bedienung und alle möglichen Funktionalitäten der Digitalkamera.

Typische Arbeitsaufgaben können selbstverständlich auch durch Arbeitsplatzbegehungen und Feldbeobachtung der Nutzer im identifiziert werden, da dies im konkreten Fall jedoch nicht möglich ist werden die Aufgaben nach ihrer Relevanz für die späteren Endnutzer aus der Gesamtfunktionalität der Digitalkamera ausgewählt.

Funktionsübersicht Digitalkamera Canon Digital Ixus 70 (Auszug aus Handbuch):

- Fotoaufnahme
- Videoaufnahme
- Zoom einstellen
- LCD-Display ein/ausschalten
- Zeitzone einstellen
- Moduswahl Kamera/Video/Wiedergabe
- Belichtungskorrektur einstellen
- Langzeitbelichtungsmodus an/aus
- Weißabgleich einstellen
- My Colors einstellen
- Kompression einstellen
- Messverfahren einstellen
- Bildfrequenz für Film einstellen
- Aufnahmeintervall für Film einstellen
- Aufnahmepixel für Foto einstellen
- Aufnahmepixel für Video einstellen
- Gesichtserkennung ein/ausschalten
- Digitalzoom Foto ein/ausschalten
- Digitalzoom Film ein/ausschalten
- Langzeitsynchronisation ein/ausschalten
- Rote Augen Reduktion ein/ausschalten
- Selbstauslöser einstellen
- AF-Hilfslicht ein/aus
- Rückblick einstellen
- Original speichern ein/aus
- Auto Kategorie ein/aus
- Überlagerung Gitternetz ein/aus
- Stummschaltung ein/aus
- Lautstärke einstellen
    - Lautstärke Startton einstellen
    - Lautstärke Betriebsgeräusch einstellen
    - Lautstärke Selbstauslöser einstellen
    - Lautstärke Auslöser einstellen
    - Lautstärke Wiedergabe einstellen
- LCD-Helligkeit einstellen
- Stromsparmodus ein/aus
    - Autoabschaltung ein/aus
    - Automatische Displayabschaltung ein/aus
- Dateinummerierung einstellen
- Ordner anlegen
    - Neuen Ordner anlegen

    - Automatischen Ordner anlegen
- Automatische Drehen an/aus
- Objektiv einfahren Zeit einstellen
- Sprache auswählen
- Videosystem einstellen
- Druckmethode einstellen
- Auf Werkseinstellungen zurücksetzen
- Thema einstellen
- Startbild einstellen
- Start-Ton einstellen
- Tastenton einstellen
- Selbstauslöser einstellen
- Auslösegeräusch einstellen
- Speicherkarte formatieren
- Bilder löschen
- Videos löschen
- Druckformat einstellen
- Übertragungseinstellungen festlegen
- Blitz einstellen
- …

Die Evaluation aller Arbeitsaufgaben ist zu aufwendig, daher werden repräsentativ für alle möglichen Funktionalitäten drei Funktionen zur Evaluation ausgewählt, die in ihrer praktischen Relevanz für den Endanwender als wichtig erachtet und häufig benutzt werden.

➔ Auswahl von drei Funktionen zur Evaluation:

- Fotoaufnahme
- Bilder löschen
- Videoaufnahme

6.1.4 Formulierung von Aufgabenszenarien

Um die Testaufgaben in einem realistischen Szenario darzustellen, werden den ausgewählten Testaufgaben plausible Alltagssituationen, die beim Umgang mit der Digitalkamera auftreten können, zugeordnet und formuliert.

Ein Aufgabenszenario besteht aus einer Situationsbeschreibung, zum Beispiel einer kurzen Geschichte, die dem Probanden als Einleitung zur Aufgabe dient und einer konkreten Arbeitsanweisung. Die Arbeitsanweisung selbst, beschreibt die zu erledigende Aufgabe innerhalb der Situationsbeschreibung.

Um ein erstes Erfolgserlebnis der Probanden zu gewährleisten, wird mit einer leichten Testaufgabe begonnen. Der unmittelbare Lerneffekt des Probanden bei der Aufgabenbearbeitung, kann durch Steigerung des Schwierigkeitsgrades der Testaufgaben berücksichtigt werden.

Aufgabenszenario 1:

Sie haben zu Weihnachten von ihrer Familie eine Digitalkamera geschenkt bekommen. Da ihre ganze Familie anwesend ist, wollen Sie die Gelegenheit gleich nutzen und ein Gruppenfoto vom Weihnachtsabend machen.

Schalten Sie die Kamera ein und nehmen Sie ein Foto auf.

Aufgabenszenario 2:

Als Sie das Gruppenfoto ihrer Familie gemacht haben, haben Sie bemerkt, dass Sie ihre Oma vergessen haben. Sie wollen die Aufnahme daher wiederholen, allerdings das vorher aufgenommene Gruppenfoto löschen.

Löschen Sie ein Foto.

Aufgabenszenario 3:

Sie gehen auf ein Konzert und möchten einen Videomitschnitt Ihres Lieblingssongs mit der Videofunktion ihrer neuen Digitalkamera aufnehmen.

Nehmen Sie ein 30-Sekunden langes Video auf.

6.1.5 Zielgruppendefinition

Die Zielgruppendefinition dient einerseits der Unterstützung der Auswahl von Probanden beim Usability-Test und andererseits der Benutzer- und Nutzungskontextdefinition für die Expertenmethoden.

Merkmale der Zielgruppendefinition sind (siehe Kapitel 2.1.4) physische Merkmale wie Alter, Geschlecht und Vorerfahrungen und Kenntnisse der Nutzer in Bezug auf die Wissensdomäne und den Nutzungskontext der Digitalkamera.

Folgende fiktive Zielgruppendefinition wird zur Definition verwendet:

- Altersstruktur: 16-59 Jahre.
- Geschlecht: m/w.
- Vorkenntnisse: keine oder nur in geringem Maße.
- Nutzer, die eine Digitalkamera zu ersten Mal benutzen (Novizen).
- Nutzer, die bereits erste Erfahrungen mit Digitalkameras besitzen (Intermediate User).
- Nutzer, die das Handbuch der Kamera vor der ersten Benutzung der Kamera nicht lesen und nur im Notfall darauf zurückgreifen.

Beispiele für Personendefinitionen innerhalb der Zielgruppe:

Testperson 1

- Männlich, 24 Jahre alt, Student.
- Hat schon erste Erfahrungen mit Digitalkameras gemacht.
- Technikaffin, ist mit vergleichbaren Geräten wie Handys, Videokameras und MP3-Playern vertraut.

Testperson 2

- Weiblich, 56 Jahre alt, Hausfrau.
- Interessiert sich nicht für technische Details und Featurevielfalt.
- Die Kamera soll ihren Zweck erfüllen, also einfach gute Fotos machen ohne viel Einstellaufwand.

Testperson 3

- Weiblich, 16 Jahre alt, Schülerin.
- Besitzt neben ihrem Handy kein vergleichbares Gerät.
- Die Kamera soll leicht zu bedienen und selbsterklärend sein.

6.2 Durchführung des CWT

Vorbereitungsphase

Die Vorbereitungsphase dient der Definition des Input des CWT.

Annahmen über die Nutzer

Hier wird definiert wer die Kamera in Zukunft bedienen wird. Es werden Annahmen über die voraussichtliche Nutzergruppe und Zielgruppe des Produkts definiert. Die Annahmen über die Nutzer beinhalten folgende Angaben:

- Nutzergruppen (Alter, Geschlecht)
- Produktbezogene und technische Vorkenntnisse und Erfahrungen
- Eigenschaften der Zielgruppe (Novizen, Hobbyfotografen oder Experten,)

Es wird die Zielgruppendefinition aus Abschnitt 6.1.4 verwendet.

Auswahl der zu analysierenden Aufgaben

Hier wird definiert, welche Aufgaben innerhalb des CWT durch die Experten bearbeitet werden. Die Aufgabenauswahl kann anhand einer Aufgabenanalyse des Produkts erfolgen und es können Beispielaufgaben aus den Kernfunktionen der Kamera abgeleitet werden. Es sollten Aufgaben ausgewählt werden, die

- repräsentativ für die Funktionalität des Produkts sind
- häufig benutzt werden, bei der späteren Nutzung durch die User
- realistische und kritisch innerhalb der Anwendung des Produkts Digitalkamera sind

Es werden die Aufgabenszenarien aus Kapitel 6.1.3 verwendet.

Festlegung der idealen Handlungsabfolge für jede Aufgabe

Hier wird definiert, wie die ideale Handlungsabfolge zur Lösung der ausgewählten Aufgaben aussieht. Diese Aktionsabfolge von Teilschritten zur Lösung einer Gesamtaufgabe wird dokumentiert und später bei der Durchführung des CWT als idealer Lösungspfad für die zu bearbeitende Aufgabe betrachtet.

Beispiel für Aufgabe 1 “Nehmen Sie ein Foto auf“:

1. Auf Knopf ON/OFF drücken um die Kamera einzuschalten
2. Auf Auslöseknopf drücken und zwei Sekunden halten, um ein Bild aufzunehmen.

Definition der Schnittstellen

Hier wird definiert, was der Nutzer von dem Interface der Kamera in jeden einzelnen Handlungsschritt zu sehen bekommt. Die Beschreibung beinhaltet die detaillierte Beschreibung der zur Verfügung stehenden Bedienelemente im jeweiligen Handlungsschritt sowie der Reaktionen und des Feedbacks des Systems für jede Aktion innerhalb einer Handlungssequenz.

Beispiel für eine Beschreibung des Systemzustandes Digitalkamera nach Betätigen des Knopfes ON/OFF:

- Das Objektiv der Kamera fährt aus
- Der LCD Bildschirm der Kamera wird aktiviert.
- Es ertönt ein kurzes Audiosignal nach Betätigung des Einschaltknopfes.

Analysephase

Die Experten bearbeiten nun die zuvor definierten Aufgaben gemeinsam Schritt für Schritt und formulieren sogenannte Erfolgs- oder Misserfolgsstories auf Basis der zuvor definierten Annahmen über den Nutzer. Dies geschieht durch die Beantwortung von vier Leitfragen innerhalb jedes Handlungsschrittes bei der Aufgabenbearbeitung.

1. Versucht der Nutzer den richtigen Effekt zu erzielen?
2. Erkennt er, dass die korrekte Aktion zur Verfügung steht?
3. Assoziiert er die korrekte Aktion mit dem gewünschten Effekt?
4. Erhält er ausreichend Feedback, um zu erkennen, dass er durch die korrekte Aktion der Lösung des Problems näher gekommen ist?

Die Beantwortung der vier Leitfragen wird detailliert begründet:

Beispiel einer Erfolgsstory der Aktionssequenz “Kamera einschalten“ innerhalb des Aufgabenszenarios “Nehmen Sie ein Foto auf“:

Auf Knopf ON/OFF drücken um Kamera einzuschalten

1. **Erfolgsstory:**

   Die Bezeichnung ON/OFF des Ein/Ausschalters der Kamera ist weit verbreitet unter vergleichbaren Elektrogeräten und sollte dem Nutzer aus früherer Erfahrung bekannt sein.

2. **Erfolgsstory:**

   Der Einschaltknopf der Kamera ist auf der Oberseite der Kamera deutlich platziert und durch die Beschriftung ON/OFF als solcher erkennbar.

3. **Erfolgsstory:**

Es steht nur ein Knopf mit der Bezeichnung ON/OFF zur Verfügung, ein Missverständnis zwischen der Intention des Nutzers (Kamera einschalten) und der erforderlichen Handlung kann ausgeschlossen werden.

4. **Erfolgsstory:**

   Nach Betätigung des Knopfes ON/OFF wird das LCD-Display der Kamera aktiviert, das Objektiv der Kamera fährt aus und er wird ein kurzes Audiosignal ausgegeben. Die Systemreaktion ist ausreichend.

Insgesamt wird der Teilschritt “Kamera einschalten“ als **Erfolgsstory** bewertet.

Eine Misserfolgsstory liegt bereits bei der Verneinung nur einer der vier Leitfragen vor. Wird eine Misserfolgsstory durch die Experten formuliert, so wird im nächsten Teilschritt der Handlungssequenz “Foto aufnehmen“ die vorherige Aktion formal als Erfolgsstory betrachtet. Das bedeutet, dass der Nutzer theoretisch zwar diesen nächsten Handlungsschritt nicht erreicht hat, praktisch aber, das die weitere Handlungsabfolge durch die Experten untersucht werden kann.

Während der Durchführung des CWT können die verschiedenen Teilschritte der Aufgabenbearbeitung über Erfassung des Displayinhaltes zur Erläuterung der Erfolgs- und Misserfolgsstories nützlich sein. Dadurch entfällt die aufwändige Beschreibung des Systemzustandes der Digitalkamera innerhalb jedes Aufgabenschrittes.

Bild 6.3 zeigt exemplarisch den Dialog “Bild löschen bestätigen“ innerhalb des Aufgabenszenarios 2. Diese Art der Zustandsbeschreibung der Kamera erleichtert den Evaluations- und Dokumentationsaufwand für die Experten erheblich, da nicht jedes Menü oder jeder Dialog vollständig verbal beschrieben werden muss.

Bild 6.3: Systemzustand “Bild löschen“ der Digitalkamera

Auf Seite 89 findet sich ein Beispielbogen für die Beantwortung der vier Leitfragen durch die Experten bei Bearbeitung des Teilschrittes 1, Kamera einschalten, innerhalb des Aufgabenszenarios 1, “Foto aufnehmen“.

Dieser Bogen enthält die wichtigsten Informationen zum untersuchten Lösungsweg, wie z.B. die Nummer des jeweiligen Aufgabenszenarios und den idealen Handlungsweg, sowie den Teilschritt, in dem sich die Experten befinden.

Cognitive Walkthrough – Digitalkamera - Arbeitsbogen Muster

Produkt: Canon Digital Ixus 70

Aufgabenszenario: 1 – Nehmen Sie ein beliebiges Foto auf

Idealer Handlungsweg

1.Schritt: Auf Einschaltknopf ON/OFF drücken

2. Schritt: Auf Auslöseknopf drücken und mindestens eine Sekunde halten.

1.Schritt: Auf Einschaltknopf ON/OFF drücken

☑ Versucht der Nutzer den richtigen Effekt zu erzielen?

Erfolgsstory: Die Bezeichnung ON/OFF des Ein/Ausschalters der Kamera ist weit verbreitet unter vergleichbaren Elektrogeräten und sollte dem Nutzer aus früherer Erfahrung bekannt sein.

☑ Erkennt er, dass die korrekte Aktion zur Verfügung steht?

Erfolgsstory: Der Einschaltknopf der Kamera ist auf der Oberseite der Kamera deutlich platziert und durch die Beschriftung ON/OFF als solcher erkennbar.

☑ Assoziiert er die korrekte Aktion mit dem gewünschten Effekt?

Erfolgsstory: Es steht nur ein Knopf mit der Bezeichnung ON/OFF zur Verfügung, ein Missverständnis zwischen der Intention des Nutzers (Kamera einschalten) und der erforderlichen Handlung kann ausgeschlossen werden.

☑ Erhält er ausreichend Feedback, um zu erkennen, dass er durch die korrekte Aktion der Lösung des Problems näher gekommen ist?

Erfolgsstory: Nach Betätigung des Knopfes ON/OFF wird das LCD-Display der Kamera aktiviert, das Objektiv der Kamera fährt automatisch aus und es wird ein kurzes Audiosignal ausgeben. Die Systemreaktion ist ausreichend.

6.3 Durchführung der Heuristischen Evaluation

Gemäß Kapitel 4.3.2 und den formulierten Aufgabenszenarien und Zielgruppendefinitionen in Kapitel 6 wird hier exemplarisch der Ablauf einer Heuristischen Evaluation am Beispiel der Digitalkamera erläutert. Als Heuristiken werden für dieses Beispiel die 12 alternativen, generellen Heuristiken von Sarodnick verwendet.

Vorevaluatives Training

Die Evaluation beginnt mit einer Einführung der Experten in die Systematik der Methode und die zu benutzenden Heuristiken. Es folgt eine grobe Einführung in das Bedienkonzept der Digitalkamera durch den Versuchsleiter. Es werden die verschiedenen Funktionalitäten, wie Foto und Videomodi erläutert und der voraussichtliche Einsatzbereich und Nutzungskontext der Kamera vorgestellt. Anschließend erfolgt die Vorstellung der Zielgruppe und der Aufgabenszenarien zur Aufgabenbearbeitung. Den Schluss des vorevaluativen Trainings bildet die Einweisung der Experten in die Dokumentation der gefundenen Probleme.

Durchführung

Die Experten bearbeiten die Aufgabeszenarien unabhängig voneinander und notieren die gefundenen Usability-Probleme bei der Aufgabenbearbeitung. Es werden zwei Evaluationsdurchgänge durchgeführt. Die Berichtsbögen enthalten den Fundort des Problems, dessen Beschreibung und die Heuristik, die dem identifizierten Problem zugrunde liegt.

Beispielbeschreibung eines Problems:

Beschreibung	Gestaltung des Einschaltknopfes ON/OFF zu klein
Fundort	Oberseite der Kamera – Bedienknopf ON/OFF
Heuristik	1 – Aufgabenangemessenheit

Tabelle 6.1: Beispielproblembeschreibung einer Heuristischen Evaluation

Auswertungsphase

Der Versuchsleiter sammelt die Berichtsbögen ein und erstellt eine Gesamtproblemliste. Anschließend bewerten die Experten gemeinsam die gefundenen Usability-Probleme der Gesamtproblemliste in ihrer Schwere von null bis vier. Das Ergebnis der Auswertungsphase ist die priorisierte Gesamtproblemliste. Tabelle 6.2 auf Seite 92 zeigt ein theoretisches Beispielergebnis einer priorisierten Problemliste bei der Evaluation der Digitalkamera. Es werden die einzelnen Probleme mit Fundort und Benennung sowie,

die den Problemen zugeordnete Heuristik aufgeführt. Anhand von diesem Beispiel erkennt man den wesentlichsten Vorteil der Heuristischen Evaluation. Durch die gemeinsame Gewichtung der identifizierten Probleme durch alle Experten, wird deren Schwere und Ausmaß deutlich.

In unten aufgeführtem Beispiel wird das Problem, der nicht vorhandenen Akkuanzeige der Heuristik Nummer drei, Selbstbeschreibungsfähigkeit, zugeordnet. Die Mittelung der einzelnen Gewichtungen der Evaluatoren ergibt eine Schwere des identifizierten Problems von 3,75. Dieses Problem sollte sofort bzw. als erstes behoben werden, da es eine hohe Priorität besitzt.

Nr.	Beschreibung	Fundort	Heu-ristik	Gefunden durch				Schwere				Mittel
				Ev1	Ev2	Ev3	Ev4	Ev1	Ev2	Ev3	Ev4	
1	Gestaltung des Einschalt-knopfes ON/OFF zu klein.	Oberseite der Kamera – Bedienknopf ON/OFF	1		X			2	2	3	2	**2,25**
2	Bezeichnung des Knopfes FUNC/SET unklar.	Rückseite der Kamera – Bedienknopf FUNC/SET	3	X		X		2	1	2	2	**1,75**
3	Akkuanzeige nicht sich-tbar im Display.	Display	3	X	X	X	X	3	4	4	4	**3,75**
4	Anzahl der Symbole im Display unübersichtlich hoch.	Display	8		X		X	1	1	1	2	**1,25**
5	Symbole im Display teil-weise nicht interpretierbar	Display	3	X	X			2	2	3	2	**2,75**

Tabelle 6.2: Beispiel einer priorisierten Problemliste

6.4 Durchführung eines Benutzertests

Begrüßung und Einführung

Nach der Begrüßung der Probanden gibt der Testleiter eine Einführung in die Digitalkamera. Innerhalb der Einführung wird der Einsatzbereich der Digitalkamera und ihre Funktionalität mit einfachen, verständlichen Worten erklärt. Desweiteren muss betont werden, dass die Probanden dass Produkt testen und nicht die Probanden getestet werden. Es wird der Testablauf erklärt und auf die Aufgabenszenarien zur Bearbeitung hingewiesen. Der Testleiter beantwortet anschließend alle offenen Fragen der Probanden zum Testablauf oder der weiteren Datenverwendung. Im Idealfall wird durch den Testleiter ein Probelauf des Tests durchgeführt um eventuelle organisatorische Probleme und Unklarheiten von vorneherein zu vermeiden.

Bei der Einführung eines Benutzer-Tests durch den Testleiter sollten allgemein folgende Punkte /20/ beachtet werden:

- Durch eine gewissenhafte Vorbereitung sollte ein reibungsloser Ablauf des Tests gewährleistet werden.
- Das Ziel des Tests ist das Produkt, also die Kamera, zu testen und nicht den Probanden.
- Da die Testteilnahme freiwillig ist, besteht jederzeit die Möglichkeit den Test abzubrechen.
- Es sind Hinweise auf jegliche Erhebungsmethoden und Aufzeichnungsgeräte wie Mikrofone und Kameras zu geben.
- Die Anonymisierung der Ergebnisse wird zugesichert.
- Vor dem Test sind alle offenen Fragen zu klären.

Aufgabenbearbeitung

Nach Beendigung der Einführungsphase bearbeitet der Proband die Aufgabenszenarien unter Beobachtung des Testleiters. Die Aufgabenszenarien steigen im Schwierigkeitsgrad um dem Probanden einen „leichten“ Einstieg in die Bedienung der Kamera zu ermöglichen und gleichzeitig dem unvermeidlichen Lerneffekt während der fortschreitenden Aufgabenbearbeitung zu berücksichtigen. Der Testleiter notiert während der Durchführung kritische Ereignisse und die benötigte Zeit der Aufgabenbearbeitung des Probanden. Die wichtigsten Informationen, die der Testleiter notieren sollte, sind die Zeit der Aufgabenbearbeitung, den Code der Testperson sowie die aufgetretenen kritischen Ereignisse während der Aufgabenbearbeitung. Das Erstellen eines Testprotokolls während des Tests ist eine einfache und effiziente Möglichkeit, die Probandeninteraktion einheitlich zu dokumentieren, siehe hierzu Tabelle 6.3 und 6.4 auf Seite 94.

Zeit	Ereignis	Bemerkung
00:00:00	TP1 betritt Raum	
00:00:12	Orientierung	
00:01:12	B TP1 A1	
00:01:55	„Der Knopf ist ja viel zu klein“	bei drücken des Einschaltknopfes
00:02:12	„Hab ich jetzt schon ein Foto gemacht oder nicht? “	Nach mehrmaligen Klicken des Auslöse-knopfes, jedoch noch kein Foto gemacht
00:02:30	„Ahhh ich muss lang draufd-rücken“	Foto gemacht
00:04:30	E TP1 A1	
00:05:40	B TP1 A2	

Tabelle 6.3: Beispielprotokoll eines Benutzertests

Legende	Erläuterung
TP X	Testperson X
B	Beginn
E	Ende
A X	Aufgabenszenario X
“…“	Freier Kommentar der Testperson wäh-rend der Aufgabenbearbeitung

Tabelle 6.4: Legende zum Beispielprotokoll eines Benutzertests

Generell sollten während der Durchführung eines Benutzer-Tests folgende Punkte /20/ beachtet werden:

- Durch ein leichtes Einstiegsszenario wird dem Probanden ein frühes Erfolgser-lebnis ermöglicht.
- Jedes Aufgabenszenario wird separat und nacheinander an den Probanden aus-gehändigt.
- Störungsfreiheit während der Durchführung des Tests.

- Der Testleiter darf beim Probanden nie den Eindruck erwecken, dass er die Aufgaben zu langsam bearbeitet.
- Minimierung der Beobachter beim Test.
- Im Falle eines betriebsinternen Tests, sollten keine direkten Vorgesetzen der Probanden anwesend sein.
- Wenn die Probanden sich deutlich unwohl fühlen, sollte der Testleiter den Test abbrechen.

Fragebogen und Nachbefragung

Nach der Aufgabenbearbeitung durch den Probanden wird dem Probanden ein Fragebogen ausgehändigt. Da die Attraktivität und praktische Qualität der Kamera beurteilt werden soll, kann der im Kapitel 4.2.10.5 beschrieben AttrakDiff Fragebogen verwendet werden.

Die Nachbefragung durch den Testleiter bildet den Abschluss des Tests. Die Nachbefragung kann als halbstandardisiertes Interview durchgeführt werden.

In einem ersten Befragungsteil können Unklarheiten, die der Testleiter während der Aufgabenbearbeitung beobachtet hat, durch gezieltes Nachfragen bei der Testperson beantwortet werden. Im zweiten Teil werden standardisierte Fragen zum getesteten Produkt gestellt.

Hier einige Beispiele für standardisierte Fragen zur Benutzung der Digitalkamera nach einem Benutzertest:

- Welchen Gesamteindruck hatten Sie von der Digitalkamera?
- Was hat Ihnen an der Digitalkamera besonders gut gefallen?
- Was hat Ihnen an der Digitalkamera weniger gut gefallen?
- Haben Sie Verbesserungsvorschläge?

6.5 Vorteile und Nachteile der verschiedenen Methoden

Im vorangegangenen fiktiven Praxisbeispiel wurde der unterschiedliche Output der Methoden am Beispiel der Bewertung einer Digitalkamera kurz erläutert und die verschiedenen Möglichkeiten der Dokumentation beschrieben. Abschließend werden in Tabelle 6.5 die Vor- und Nachteile der einzelnen Methoden stichpunktartig erläutert:

Verfahren	Vorteile	Nachteile
Cognitive Walkthrough	+ Einziges Verfahren zur Überprüfung der Erlernbarkeit des Bedienprinzips der Digitalkamera ohne Probanden. + Detaillierte Begründungen der Interaktion durch die Formulierung der Erfolgs- und Misserfolgsstories. + Ableitung von Lösungsvorschlägen und Optimierungshinweisen aus der Negation der vier Leitfragen. + minimaler technischer Aufwand. + Standardisierte Vorgehensweise. + Einheitliche Dokumentation	- Zeitaufwändig, da für jede Interaktion ein separater CWT durchgeführt werden muss - Keine Erhebung von qualitativen und quantitativen Nutzerdaten möglich. - Expertenbasiertes Verfahren, dadurch begrenzte Vorhersagekraft.
Heuristische Evaluation	+ Umfassende Bewertung der Gesamtgestaltung der Digitalkamera. + Gewichtete Problemliste zeigt Probleme unterschiedlicher Priorität. + Minimaler technischer Aufwand. + Flexibel und schnell anzuwenden.	- Keine Erhebung von qualitativen und quantitativen Nutzerdaten möglich. - Expertenbasiertes Verfahren, dadurch begrenzte Vorhersagekraft.

Tabelle 6.5: Übersicht der Vor- und Nachteile der Verfahren

Verfahren	Vorteile	Nachteile
Benutzertest	+ Höchste Vorhersagekraft durch Verwendung von realen Probanden. + Erhebung von objektiven Daten wie Fehlerraten und Dauer der Aufgabenbearbeitung. + Messung der Akzeptanz und Zufriedenheit durch Interviews.	- Aufwändig zu realisieren. - Keine einheitliche Form der Dokumentation und Auswertung des Tests vorhanden.
Fragebogen	+ Höchste Vorhersagekraft durch Verwendung von realen Probanden. + Standardisierte Befragung und Auswertung der Probandenmeinung zur Attraktivität der Digitalkamera. + Einfach, schnell und kostengünstig zu realisieren.	- Keine Identifikation von konkreten Usability-Problemen, da nur generelle Problembereiche aufgezeigt werden. - Anwendung nur bei einer ausreichenden Zahl von Probanden sinnvoll.

Tabelle 6.5: Übersicht der Vor- und Nachteile der Verfahren - Fortsetzung

7 Fazit und Kommentar

In den vorangegangen Kapiteln wurden insgesamt fünf verschiedene Methoden zur Evaluation von Produkten und interaktiven Systemen vorgestellt. Anhand eines theoretischen Anwendungsbeispiels, der Bewertung einer Digitalkamera, wurde der praktische Ablauf der Methoden erläutert.

Es ist offensichtlich, dass sich die Methoden für verschiedene Zielsetzungen und Evaluationsanliegen eignen.

Das Ziel einer heuristischen Evaluation, ist die Identifikation von Verstößen gegen festdefinierte Usability-Prinzipien. Der CWT versucht den Problemlöseprozess eines Benutzers zu modellieren und darauf aufbauend, das Interaktionsverhalten des Benutzers vorherzusagen. KLM-GOMS, ermöglicht die Bewertung der Effizienz von Gestaltungsalternativen in Bildschirmsystemen.

Der Benutzertest, als das klassische, empirische Verfahren zur Bewertung der Gebrauchstauglichkeit eines Produktes, ist in seiner Zielsetzung sehr variabel. Abhängig von dem Ziel des Tests und den Anforderungen der Auftraggeber, können passende Erhebungsmethoden zur Bewertung ausgewählt und angewendet werden. Außerdem ist er für eine ganze Reihe von Zielsetzungen, wie beispielsweise Performance-Messungen und Akzeptanztests unersetzbar. Fragebögen wiederrum, eignen sich zur einfachen und kostengünstigen Erhebung von subjektiven Benutzermeinungen in großer Zahl. Ihre Auswertung erfolgt nach den Prinzipien der Statistik. Über Mittelung und Standardabweichung können Problembereiche in einem interaktiven System aufgezeigt werden.

Die Auswahl einer geeigneten Methode bzw. des entsprechenden Erhebungsverfahrens hängt primär vom Evaluationsziel, also der Frage was gemessen oder bewertet werden soll, ab. Rahmenbedingungen, wie die Verfügbarkeit geeigneter Testpersonen oder der Entwicklungsstatus des interaktiven Systems sowie Zeit und Kosten müssen bei der Auswahl einer geeigneten Methode beachtet werden.

Eine heuristische Evaluation erlaubt zwar eine umfassende Bewertung der Gebrauchstauglichkeit eines Produktes durch Experten, kann jedoch nicht die Zufriedenheit der Benutzer oder Fehlerhäufigkeit bei der Interaktion mit dem Produkt beurteilen. Der Fragebogen ISONORM 9241/10 dient der Überprüfung der Normenkonformität mit ISO 9241, AttrakDiff hingegen bewertet letztlich die Attraktivität eines Produktes und QUIS misst die Zufriedenheit der Nutzer beim Umgang mit einer Software. Der CWT, die Heuristische Evaluation und KLM-GOMS, also die expertenbasierten Methoden, können bereits in frühen Stadien des Produktentwicklungsprozesses eingesetzt werden, wenn noch keine Benutzertests durchgeführt werden können. Dadurch können nichtfunktionale Papierprototypen oder Designentwürfe eines Interfaces bereits einer ersten Überprü-

fung unterzogen werden. Das hat den Vorteil, dass Änderungen zu diesem Zeitpunkt der Entwicklung wesentlich kostengünstiger zu realisieren sind. Die Annahme, dass die künftigen Benutzer das Produkt auch wirklich akzeptieren und keine Probleme bei der Bedienung haben werden kann, in letzter Konsequenz, jedoch nur durch empirische Verfahren überprüft werden. Auch die Erfassung quantitativer Daten, wie der Zeit für eine Aufgabenbearbeitung, oder die Anzahl an Fehlern während einer Aufgabenbearbeitung, ist nur durch einen Benutzertest realisierbar.

Innerhalb eines gesamten Produktentwicklungszyklus ist eine Kombination verschiedener Methoden sinnvoll und empfehlenswert. Das Interface eines Produktes muss nicht funktionsfähig und marktreif sein um einer Bewertung durch die Expertenmethoden unterzogen zu werden. Bereits im Konzeptstadium ist die Bewertung eines Interfacelayouts durch die Expertenmethoden möglich. Sie eignen sich für die Bewertung und Umgestaltung nicht-funktionaler Papierprototypen oder Interfacelayouts.

Sobald ein funktionaler Prototyp zur Verfügung steht, können Benutzertests durchgeführt werden. Durch die Benutzertests werden Ergebnisse mit hoher Vorhersagekraft erzielt und eventuelle Probleme erkannt, die innerhalb der Expertenmethoden unentdeckt geblieben sind.

Da sich Experten über die Effektivität und Effizienz eines Produkts irren können, sind stets die Benutzer nach ihrem subjektiven Eindruck über Effektivität und Effizienz zu befragen. Wenn Experten und Benutzer im Urteil abweichen, so haben bei mangelnder Zufriedenstellung immer die Benutzer Recht /3/.

Die im Rahmen dieser Studienarbeit vorgestellten Methoden stellen eine Auswahl der in der Praxis und Literatur am weitesten verbreiteten Verfahren zur Usability-Evaluation dar. Anhand eines fiktiven Praxisbeispiels wurde versucht, die tendenziellen Vor- und Nachteile, sowie den jeder Methode zugrunde liegende Ansatz, gegeneinander abzuwägen und vorzustellen.

Die vorgestellte Bewertung, die anhand gemeinsamer Kriterien, die verschiedenen Ausprägungen der einzelnen Methoden zur Usability-Evaluation quantifiziert, erlaubt eine erste grobe Einschätzung der Unterschiede der einzelnen Erhebungsverfahren. Diese Bewertung basiert auf einer Literaturrecherche und theoretischen Überlegungen zum Ansatz und der Durchführung der jeweiligen Methoden.

Die Menschengerechte, ergonomische Gestaltung von technischen Produkten und interaktiven Systemen wird in naher Zukunft aufgrund des unaufhaltsamen, technologischen Fortschritts und der zunehmenden Multifunktionalisierung, ein immer wichtiger Faktor für erfolgreiche, konkurrenzfähige Produkte werden. Die nutzerzentrierte Gestaltung eines technischen Produktes und die Bewertung desselbigen, mithilfe der vorgestellten Usability-Evaluationsmethodik sind daher die entscheidenden Werkzeuge, um Produkte mit einer hohen ergonomischen Qualität herstellen zu können.

8 Quellenverzeichnisse

8.1 Literatur- und Quellenverzeichnis

/1/ /Bortz_02/ Bortz J., Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, Berlin, Springer Verlag, 2002.

/2/ /Gesetz_06/ Bundesministerium der Justiz: Verordnung über Sicherheit und Gesundheitsschutz bei der Arbeit an Bildschirmgeräten, http://bundesrecht.juris.de/bundesrecht/bildscharbv/gesamt.pdf geprüft am 01.05.08, 2006.

/3/ /DATECH_08/ Deutsche Akkreditierungsstelle Technik in der TGA GmbH, Leitfaden Usability v.1.1, http://www.datech.de/share/files/Leitfaden-Usability.pdf geprüft am 01.05.08, 2008.

/4/ /DIN_66050_80/ DIN 66050: Gebrauchstauglichkeit – Begriff, Berlin, Beuth Verlag, 1980.

/5/ /DIN_13407_00/ DIN EN ISO 13407: Benutzer-orientierte Gestaltung interaktiver Systeme, Berlin, Beuth Verlag, 2000.

/6/ /DIN_14915-1_03/ DIN EN ISO 14915-1: Software-Ergonomie für Multimedia-Benutzungsschnittstellen, Teil 1: Gestaltungsgrundsätze und Rahmenbedingungen, Berlin, Beuth Verlag, 2003, S.10.

/7/ /DIN_9241-11_99/ DIN EN ISO 9241-11: Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 11: Anforderungen an die Gebrauchstauglichkeit – Leitsätze, Berlin, Beuth Verlag ,1999.

/8/ /DIN_9241-110_06/ DIN EN ISO 9241-110: Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung, Berlin, Beuth Verlag, 2006.

/9/ /Gediga_02/ Evaluation in der Software-Ergonomie: Methoden und Modelle im Software-Entwicklungszyklus – Zeitschrift für Psychologie Nr.210 (1), 2002, S.40-57.

/10/ /Hassenzahl_03/ Hassenzahl M., Mensch und Computer 2003 – Interaktion in Bewegung – Attrakdiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität, Stuttgart, 2003, S.187-196.

/11/ /Hilbert_00/ Hilbert D., Extracting Usability Information from User Interface Events – ACM Computing Surveys 32 – Number 4, ACM, 2000, S.384-421.

/12/ /Jeffries_91/ Jeffries R., User interface evaluation in the real world: a comparison of four techniques - Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology, ACM, 1991.

/13/ /Klemmert_01/ Klemmert H., Erweiterung der Heuristischen Evaluation für

sprachgesteuerte Softwaresysteme – Ergebnisse der Evaluation, Hess, W. & Stöber, K. (Hrsg.), Elektronische Sprachsignalverarbeitung 12. Konferenz, Studientexte zur Sprachkommunikation (Bd. 23, S. 53-60), w.e.b. Universitätsverlag., 2001.

/14/ /Kranz_97/ Kranz T., Einführung in die klassische Testtheorie, Eschborn, Klotz Verlag, 1997, S.5.

/15/ /Landauer_93/ Landauer T. & Nielsen J., A Mathematical Model of the Finding of Usability Problems - Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems, 1993, S.206-213.

/16/ /Lavery_97/ Lavery D., Comparison of evaluation methods using structured usability problem reports – Behaviour and Information Technology, 1997, S.246-266.

/17/ /Lewis_90/ Lewis C., Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces – ACM CHI 1990 Procedings, ACM, 1990.

/18/ /Mackay_89/ Mackay W., Introduction to this special issue on video as a research and design tool – ACM SIGCHI Bulletin 2 - Nummer 2, 1989, S.48-50.

/19/ /Maier_07/ Maier T., Technisches Design I/II – Manuskript zur Vorlesung, Universität Stuttgart, 2007.

/20/ /Nielsen_93/ Nielsen J., Usability Engineering, San Francisco, Academic Press Professional, 1993.

/21/ /Nielsen_93a/ Nielsen J., Usability Engineering, San Francisco, Academic Press Professional, 1993, S.202.

/22/ /Nielsen_93b/ Nielsen J., Usability Engineering, San Francisco, Academic Press Professional, 1993, S.195.

/23/ /Nielsen_93c/ Nielsen J., Usability Engineering, San Francisco, Academic Press Professional, 1993, S.203.

/24/ /Nielsen_94/ Nielsen J., Usability Inspection Methods, New York, Wiley John + Sons, 1994.

/25/ /Nielsen_94a/ Nielsen J., Usability Inspection Methods, New York, Wiley John + Sons, 1994, S.1.

/26/ /Nielsen_94b/ Nielsen J., Usability Inspection Methods, New York, Wiley John + Sons, 1994, S. S.29-30.

/27/ /Nielsen_94c/ Nielsen J., Usability Inspection Methods, New York, New York, Wiley John + Sons, 1994, S.27.

/28/ /Polson_92/ Polson P., Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces – International Journal of Man-Machine Studies 36, Polson, ACM, 1992, S.741-773.

/29/ /Polson_90/ Polson P., Theory-Based Design for Easily Learned Interfaces – HCI Number 5, ACM, 1990, S.191-220.

/30/ /Prümper_93/ Prümper J., Die Evaluation von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm ISO 9241 Teil 10 als Beitrag zur partizipativen Systemgestaltung - ein Fallbeispiel, Software-Ergonomie '93: Von der Benutzungsoberfläche zur Arbeitsgestaltung, Stuttgart, Teubner Verlag, 1993, S.145-156.

/31/ /Rieman_95/ Rieman J., Usability Evaluation with the Cognitive Walkthrough http://www.sigchi.org/chi95/proceedings/tutors/jr_bdy.htm geprüft am 01.05.08.

/32/ /Sarodnick_06/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006.

/33/ /Sarodnick_06a/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006, S.22.

/34/ /Sarodnick_06b/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006, S.171.

/35/ /Sarodnick_06c/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006, S.140-141.

/36/ /Sarodnick_06d/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006, S.147.

/37/ /Sarodnick_06e/ Sarodnick F., Methoden der Usability Evaluation – Wissenschaftliche Grundlagen und praktische Anwendung, Bern, Verlag Hans Huber, 2006, S.185-190.

/38/ /Schmidts_07/ Schmidts H., Usability-Evaluation: Eine Studie zur Identifizierung von Nutzungsproblemen mittels Eye-Tracking-Parametern, Saarbrücken, Vdm Verlag, 2007.

/39/ /Shneiderman_02/ Shneiderman B., User Interface Design, Frechen, mitp, 2002.

/40/ /Spencer_00/ Spencer R., The Streamlined Cognitive Walkthrough Method, Working Around Social Constraints Encountered in a Software Development Company - CHI, Spencer, ACM, 2000.

/41/ /NET_1/ http://www.techsmith.com/purchase/default.asp#morae geprüft am 01.05.08.

/42/ /NET_2/ http://www.useit.com/alertbox/20000319.html geprüft am 01.05.08.

/43/ /NET_3/ http://techfee.washington.edu/proposals/2008-052-1/ geprüft am 01.05.08.

/44/ /NET_4/ http://www.attrakdiff.de/ geprüft am 01.05.08.

/45/ /NET_5/ http://www.ergoonline.de/site.aspx?url=html/software/verfa

hren _zur_beurteilung der/fragebogen_isonorm_online.htm geprüft am 01.05.08.

/46/ /NET_6/ http://lap.umd.edu/quis/ geprüft am 01.05.08.

/47/ /NET_7/ http://phillips.rmc.ca/courses/459-2007/lectures/05-goms-klm.html geprüft am 01.05.08.

/48/ /NET_8/ http://www.canon.de/For_Home/Product_Finder/Cameras/Digital_Camera/IXUS/Digital_IXUS_70/index.asp geprüft am 01.05.08.

/49/ /NET_9/ http://www3.canon.de/images/pro/fot/kom/ixu/file/DIGITAL_IXUS_70_SILBER_MANUAL.zip geprüft am 01.05.08.

/50/ /BIL_1/ DIN EN ISO 9241-11: Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 11: Anforderungen an die Gebrauchstauglichkeit – Leitsätze, Beuth Verlag ,1999, S.6.

/51/ /BIL_2/ DIN EN ISO 9241-110: Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung, Beuth Verlag, 2006, S.22.

/52/ /BIL_4/ Nielsen J., Usability Engineering, San Francisco, Academic Press Professional, 1993., S.201.

/53/ /BIL_5/ http://www.noldus.com/site/content/images/usab%20labs/icts_3.jpg geprüft am 01.05.08.

/54/ /BIL_6/ http://www.rerc-ami.org/ami/projects/d/1/manual/suitcase.jpg geprüft am 01.05.08.

/55/ /BIL_7/ http://www.useit.com/alertbox/20000319_problemfindingcurve.gif geprüft am 01.05.08.

/56/ /BIL_8/ http://www.ixbt.com/short/images/x120.jpg geprüft am 01.05.08.

/57/ /BIL_9/ http://stephenslighthouse.sirsidynix.com/archives/2006/07/eye_tracking_an.html geprüft am 01.05.08.

/58/ /BIL_10/ http://www.attrakdiff.de/files/wissenschaft_de.gif geprüft am 01.05.08.

/59/ /BIL_11/ http://www.attrakdiff.de/files/ab_de.gif geprüft am 01.05.08.

/60/ /BIL_12/ http://www.attrakdiff.de/Services/Demo/ geprüft am 01.05.08.

/61/ /BIL_13/ http://www.attrakdiff.de/Services/Demo/ geprüft am 01.05.08.

/62/ /BIL_14/ http://www.ergonomie-online.de/site.aspx?url=html/service/download_area/isonorm.doc geprüft am 01.05.08.

/63/ /BIL_15/ http://www.useit.com/papers/heuristic/heuristic_matrix.gif geprüft am 01.05.08.

/64/ /BIL_16/ http://www.useit.com/papers/heuristic/heur_eval_finding_curve.gif geprüft am 01.05.08.

/65/ /BIL_17/ Lewis C., Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces – ACM CHI

1990 Proceedings, ACM, 1990, S.237.

/66/ /BIL_18/ Polson P., Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces – International Journal of Man-Machine Studies 36, Polson, ACM, 1992, S.741-773.

/67/ /BIL_19/ http://munin.bui.haw-hamburg.de/amoll/use/beispiele-a-moll/Arbeitsbogen_Cognitive-Walkthrough.pdf geprüft am 01.05.08.

/68/ /TAB_02/ http://www.useit.com/papers/uselabs.html am 01.05.08

/69/ /TAB_03/ http://edocs.tu-berlin.de/diss/2004/backhaus_claus.pdf am 01.05.08.

8.2 Abbildungsverzeichnis

Seite

Bild 3.1: Anwendungsrahmen für Gebrauchstauglichkeit /50/ 13
Bild 3.2: Konzept Gebrauchstauglichkeit /52/ 16
Bild 4.1: Gliederung der Methoden zur Usability-Evaluation 22
Bild 4.2: Grundriss eines beispielhaften Aufbaus eines Usability-Labs /52/ 25
Bild 4.3: ICTS Usability-Lab: Kontrollraum mit Blick auf Testraum /53/ 25
Bild 4.4: Mobiles Usability Lab /54/ 28
Bild 4.5: Problemfindungskurve – Benutzertest /55/ 31
Bild 4.6: Ein Eyetracker der Firma Tobii /56/ 38
Bild 4.7: Hotspots einer Eyetrackingstudie /57/ 39
Bild 4.8: Arbeitsmodell des AttrakDiff /58/ 42
Bild 4.9: Portfolio der “Charakterbereiche“ des untersuchten Produkts /59/ 43
Bild 4.10: Diagramm der Mittelwerte /60/ 43
Bild 4.11: Verteilung der Mittelwerte auf die semantischen Differentiale /61/ 44
Bild 4.12: Fragebogen ISONORM 9241/10 – Subskala Steuerbarkeit /62/ 46
Bild 4.13: Verteilung: 19 Evaluatoren; 16 Usability Probleme /63/ 54
Bild 4.14: Problemfindungskurve Heuristische Evaluation /64/ 55
Bild 4.15: Erste CWT Anleitung für einen einzelnen Arbeitsschritt /65/ 62
Bild 4.16: CWT, zweite Version, Analysephase /28/ 63
Bild 4.17: Beispielbogen eines CWT für eine Website /67/ 66
Bild 6.1: Stellteile und Anzeigen – Kameraoberseite 79
Bild 6.2: Stellteile und Anzeigen – Kamerarückseite 79
Bild 6.3: Systemzustand “Bild löschen“ der Digitalkamera 88

8.3 Tabellenverzeichnis

Seite

Tabelle 2.1: Kompetenzen eines Usability-Experten 9
Tabelle 4.1: Ergebnisse einer Untersuchung von 13 Usability-Labs /68/ 27
Tabelle 4.2: Schwereskala der Heuristischen Evaluation 53
Tabelle 4.3: Beispiel für Heuristiken einer Spritzenpumpe /69/ 56
Tabelle 4.4: Operatoren des KLM-GOMS Modells 70
Tabelle 4.5: Zeit pro Anschlag im KLM-GOMS-Modell 70
Tabelle 5.1: Legende zur Bewertung 73
Tabelle 5.2: Bewertung der Methoden 76
Tabelle 6.1: Beispielproblembeschreibung einer Heuristischen Evaluation 90
Tabelle 6.2: Beispiel einer priorisierten Problemliste 92
Tabelle 6.3: Beispielprotokoll eines Benutzertests 94
Tabelle 6.4: Legende zum Beispielprotokoll eines Benutzertests 94
Tabelle 6.5: Übersicht der Vor- und Nachteile der Verfahren 96

Rechtlicher Hinweis

Die vorliegende Arbeit wurde unter Anleitung vornehmlich des betreuenden Wissenschaftlichen Mitarbeiters des Instituts, Herrn Dr.-Ing. M. Schmid, angefertigt. Sämtliche Ergebnisse wurden in stetiger Diskussion und in enger Abstimmung mit der dem Betreuer erarbeitet.

Der Verfasser dieser Arbeit stimmt einer Veröffentlichung der Ergebnisse durch das Institut zu und wird diesbezügliche eigene Veröffentlichungen nur in Abstimmung mit dem Institutsleiter und dem betreuenden Mitarbeiter vornehmen.

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich, Philipp Jordan, die vorliegende Studienarbeit selbstständig angefertigt, keine anderen als die angegebenen Hilfsmittel benutzt und sowohl wörtliche, als auch sinngemäß entlehnte Stellen als solche kenntlich gemacht habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Ort, Datum	Unterschrift